(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/79454 A1

(51) International Patent Classification[7]: C12N 5/10, 15/12, 15/63, 15/64, C07K 14/435, 14/47

(21) International Application Number: PCT/US01/11797

(22) International Filing Date: 11 April 2001 (11.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/196,603 13 April 2000 (13.04.2000) US
60/199,417 24 April 2000 (24.04.2000) US

(71) Applicants (for all designated States except US): SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US). SMITHKLINE BEECHAM P.L.C. [GB/GB]; New Horizons Court, Great West Road, Brentford, Middlesex TW8 9EP (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): AGARWAL, Pankaj [IN/US]; 251 West DeKalb Pike, King of Prussia, PA 19406 (US). MURDOCH, Paul, R. [GB/GB]; New Frontiers Science Park South, Third Avenue, Harlow, Essex CM19 5AW (GB). RIZVI, Safia, K. [PK/US]; 4617 Pine Street, Philadelphia, PA 19143 (US). SMITH, Randall, F. [US/US]; 4138 Presidential Drive, Lafayette Hill, PA 19444 (US). XIANG, Zhaoying [CN/US]; 2413 Ridgeway, Fort Lee, NJ 07024 (US).

(74) Agents: GIMMI, Edward, R. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/79454 A1

(54) Title: NOVEL COMPOUNDS

(57) Abstract: Polypeptides and polynucleotides of the genes set forth in Table I and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing polypeptides and polynucleotides of the genes set forth in Table I in diagnostic assays.

Novel Compounds

Field of Invention

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in diagnosis and in identifying compounds that may be agonists, antagonists that are potentially useful in therapy, and to production of such polypeptides and polynucleotides. The polynucleotides and polypeptides of the present invention also relate to proteins with signal sequences which allow them to be secreted extracellularly or membrane-associated (hereinafter often referred collectively as secreted proteins or secreted polypeptides).

Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces "functional genomics", that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superseding earlier approaches based on "positional cloning". A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

Proteins and polypeptides that are naturally secreted into blood, lymph and other body fluids, or secreted into the cellular membrane are of primary interest for pharmaceutical research and development. The reason for this interest is the relative ease to target protein therapeutics into their place of action (body fluids or the cellular membrane). The natural pathway for protein secretion into extracellular space is the endoplasmic reticulum in eukaryotes and the inner membrane in prokaryotes (Palade, 1975, Science, 189, 347; Milstein, Brownlee, Harrison, and Mathews, 1972, Nature New Biol., 239, 117; Blobel, and Dobberstein, 1975, J. Cell. Biol., 67, 835). On the other hand, there is no known natural pathway for exporting a protein from the exterior of the cells into the cytosol (with the exception of pinocytosis, a mechanism of snake venom toxin intrusion into cells). Therefore targeting protein therapeutics into cells poses extreme difficulties.

The secreted and membrane-associated proteins include but are not limited to all peptide hormones and their receptors (including but not limited to insulin, growth hormones, chemokines, cytokines, neuropeptides, integrins, kallikreins,

lamins, melanins, natriuretic hormones, neuropsin, neurotropins, pituitiary hormones, pleiotropins, prostaglandins, secretogranins, selectins, thromboglobulins, thymosins), the breast and colon cancer gene products, leptin, the obesity gene protein and its receptors, serum albumin, superoxide dismutase, spliceosome proteins, 7TM (transmembrane) proteins also called as G-protein coupled receptors, immunoglobulins, several families of serine proteinases (including but not limited to proteins of the blood coagulation cascade, digestive enzymes), deoxyribonuclease I, etc.

Therapeutics based on secreted or membrane-associated proteins approved by FDA or foreign agencies include but are not limited to insulin, glucagon, growth hormone, chorionic gonadotropin, follicle stimulating hormone, luteinizing hormone, calcitonin, adrenocorticotropic hormone (ACTH), vasopressin, interleukines, interferones, immunoglobulins, lactoferrin (diverse products marketed by several companies), tissue-type plasminogen activator (Alteplase by Genentech), hyaulorindase (Wydase by Wyeth-Ayerst), dornase alpha (Pulmozyme\ by Genentech), Chymodiactin (chymopapain by Knoll), alglucerase (Ceredase by Genzyme), streptokinase (Kabikinase by Pharmacia) (Streptase by Astra), etc. This indicates that secreted and membrane-associated proteins have an established, proven history as therapeutic targets. Clearly, there is a need for identification and characterization of further secreted and membrane-associated proteins which can play a role in preventing, ameliorating or correcting dysfunction or disease, including but not limited to diabetes, breast-, prostate-, colon cancer and other malignant tumors, hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities, asthma, manic depression, dementia, delirium, mental retardation, Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental or sexual development disorders, and dysfunctions of the blood cascade system including those leading to stroke. The proteins of the present invention which include the signal sequences are also useful to further elucidate the mechanism of protein transport which at present is not entirely understood, and thus can be used as research tools.

Summary of the Invention

The present invention relates to particular polypeptides and polynucleotides of the genes set forth in Table I, including recombinant materials and methods for their production. Such polypeptides and polynucleotides are of interest in relation to methods of treatment of certain diseases, including, but not limited to, the diseases set forth in Tables III and V, hereinafter referred to as "diseases of the invention". In a further aspect, the invention relates to methods for identifying agonists and antagonists (*e.g.*, inhibitors) using the

materials provided by the invention, and treating conditions associated with imbalance of polypeptides and/or polynucleotides of the genes set forth in Table I with the identified compounds. In still a further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate activity or levels the genes set forth in Table I. Another aspect of the invention concerns a polynucleotide comprising any of the nucleotide sequences set forth in the Sequence Listing and a polypeptide comprising a polypeptide encoded by the nucleotide sequence. In another aspect, the invention relates to a polypeptide comprising any of the polypeptide sequences set forth in the Sequence Listing and recombinant materials and methods for their production. Another aspect of the invention relates to methods for using such polypeptides and polynucleotides. Such uses include the treatment of diseases, abnormalities and disorders (hereinafter simply referred to as diseases) caused by abnormal expression, production, function and or metabolism of the genes of this invention, and such diseases are readily apparent by those skilled in the art from the homology to other proteins disclosed for each attached sequence. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with the imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate activity or levels of the secreted proteins of the present invention.

Description of the Invention

In a first aspect, the present invention relates to polypeptides the genes set forth in Table I. Such polypeptides include:

(a) an isolated polypeptide encoded by a polynucleotide comprising a sequence set forth in the Sequence Listing, herein when referring to polynucleotides or polypeptides of the Sequence Listing, a reference is also made to the Sequence Listing referred to in the Sequence Listing;

(b) an isolated polypeptide comprising a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(c) an isolated polypeptide comprising a polypeptide sequence set forth in the Sequence Listing;

(d) an isolated polypeptide having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(e) a polypeptide sequence set forth in the Sequence Listing; and

(f) an isolated polypeptide having or comprising a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to a polypeptide sequence set

forth in the Sequence Listing;

(g) fragments and variants of such polypeptides in (a) to (f).

Polypeptides of the present invention are believed to be members of the gene families set forth in Table II . They are therefore of therapeutic and diagnostic interest for the reasons set forth in Tables III and V. The biological properties of the polypeptides and polynucleotides of the genes set forth in Table I are hereinafter referred to as "the biological activity" of polypeptides and polynucleotides of the genes set forth in Table I. Preferably, a polypeptide of the present invention exhibits at least one biological activity of the genes set forth in Table I.

Polypeptides of the present invention also include variants of the aforementioned polypeptides, including all allelic forms and splice variants. Such polypeptides vary from the reference polypeptide by insertions, deletions, and substitutions that may be conservative or non-conservative, or any combination thereof. Particularly preferred variants are those in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acids are inserted, substituted, or deleted, in any combination.

Preferred fragments of polypeptides of the present invention include an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids from an amino acid sequence set forth in the Sequence Listing, or an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids truncated or deleted from an amino acid sequence set forth in the Sequence Listing. Preferred fragments are biologically active fragments that mediate the biological activity of polypeptides and polynucleotides of the genes set forth in Table I, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also preferred are those fragments that are antigenic or immunogenic in an animal, especially in a human.

Fragments of a polypeptide of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention. A polypeptide of the present invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence that contains secretory or leader sequences, pro-sequences, sequences that aid in purification, for instance multiple histidine residues, or an additional sequence for stability during recombinant production.

Polypeptides of the present invention can be prepared in any suitable manner, for

instance by isolation form naturally occurring sources, from genetically engineered host cells comprising expression systems (*vide infra*) or by chemical synthesis, using for instance automated peptide synthesizers, or a combination of such methods. Means for preparing such polypeptides are well understood in the art.

In a further aspect, the present invention relates to polynucleotides of the genes set forth in Table I. Such polynucleotides include:

(a) an isolated polynucleotide comprising a polynucleotide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polynucleotide sequence set forth in the Sequence Listing;

(b) an isolated polynucleotide comprising a polynucleotide set forth in the Sequence Listing;

(c) an isolated polynucleotide having at least 95%, 96%, 97%, 98%, or 99% identity to a polynucleotide set forth in the Sequence Listing;

(d) an isolated polynucleotide set forth in the Sequence Listing;

(e) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(f) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide set forth in the Sequence Listing;

(g) an isolated polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(h) an isolated polynucleotide encoding a polypeptide set forth in the Sequence Listing;

(i) an isolated polynucleotide having or comprising a polynucleotide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to a polynucleotide sequence set forth in the Sequence Listing;

(j) an isolated polynucleotide having or comprising a polynucleotide sequence encoding a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to a polypeptide sequence set forth in the Sequence Listing; and

polynucleotides that are fragments and variants of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

Preferred fragments of polynucleotides of the present invention include an isolated polynucleotide comprising an nucleotide sequence having at least 15, 30, 50 or 100 contiguous nucleotides from a sequence set forth in the Sequence Listing, or an isolated

polynucleotide comprising a sequence having at least 30, 50 or 100 contiguous nucleotides truncated or deleted from a sequence set forth in the Sequence Listing.

Preferred variants of polynucleotides of the present invention include splice variants, allelic variants, and polymorphisms, including polynucleotides having one or more single nucleotide polymorphisms (SNPs).

Polynucleotides of the present invention also include polynucleotides encoding polypeptide variants that comprise an amino acid sequence set forth in the Sequence Listing and in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acid residues are substituted, deleted or added, in any combination.

In a further aspect, the present invention provides polynucleotides that are RNA transcripts of the DNA sequences of the present invention. Accordingly, there is provided an RNA polynucleotide that:

(a) comprises an RNA transcript of the DNA sequence encoding a polypeptide set forth in the Sequence Listing;

(b) is a RNA transcript of a DNA sequence encoding a polypeptide set forth in the Sequence Listing;

(c) comprises an RNA transcript of a DNA sequence set forth in the Sequence Listing; or

(d) is a RNA transcript of a DNA sequence set forth in the Sequence Listing;

and RNA polynucleotides that are complementary thereto.

The polynucleotide sequences set forth in the Sequence Listing show homology with the polynucleotide sequences set forth in Table II. A polynucleotide sequence set forth in the Sequence Listing is a cDNA sequence that encodes a polypeptide set forth in the Sequence Listing. A polynucleotide sequence encoding a polypeptide set forth in the Sequence Listing may be identical to a polypeptide encoding a sequence set forth in the Sequence Listing or it may be a sequence other than a sequence set forth in the Sequence Listing, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes a polypeptide set forth in the Sequence Listing. A polypeptide of a sequence set forth in the Sequence Listingis related to other proteins of the gene families set forth in Table II, having homology and/or structural similarity with the polypeptides set forth in Table II. Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one activity of the genes set forth in Table I.

Polynucleotides of the present invention may be obtained using standard cloning and screening techniques from a cDNA library derived from mRNA from the tissues set forth in Table IV (see for instance, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. A polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Polynucleotides that are identical, or have sufficient identity to a polynucleotide sequence set forth in the Sequence Listing, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification reaction (for instance, PCR). Such probes and primers may be used to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from other species) that have a high sequence similarity to sequences set forth in the Sequence Listing, typically at least 95% identity. Preferred probes and primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50, if not at least 100 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

A polynucleotide encoding a polypeptide of the present invention, including homologs from other species, may be obtained by a process comprising the steps of screening a library under stringent hybridization conditions with a labeled probe having a sequence set forth in the Sequence Listing or a fragment thereof, preferably of at least 15

nucleotides; and isolating full-length cDNA and genomic clones containing the polynucleotide sequence set forth in the Sequence Listing. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes isolated polynucleotides, preferably with a nucleotide sequence of at least 100, obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence set forth in the Sequence Listing or a fragment thereof, preferably of at least 15 nucleotides.

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide does not extend all the way through to the 5' terminus. This is a consequence of reverse transcriptase, an enzyme with inherently low "processivity" (a measure of the ability of the enzyme to remain attached to the template during the polymerisation reaction), failing to complete a DNA copy of the mRNA template during first strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., Proc Nat Acad Sci USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon (trade mark) technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon (trade mark) technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adapter specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes

well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems comprising a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Polynucleotides may be introduced into host cells by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology (1986) and Sambrook *et al.*(*ibid*). Preferred methods of introducing polynucleotides into host cells include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, micro-injection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector that is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate polynucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, (*ibid*). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they

may be heterologous signals.

If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and/or purification.

Polynucleotides of the present invention may be used as diagnostic reagents, through detecting mutations in the associated gene. Detection of a mutated form of a gene is characterized by the polynucleotides set forth in the Sequence Listing in the cDNA or genomic sequence and which is associated with a dysfunction. Will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques well known in the art.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or it may be amplified enzymatically by using PCR, preferably RT-PCR, or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled nucleotide sequences of the genes set forth in Table I. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence difference may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (see, for instance, Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed

by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, Proc Natl Acad Sci USA (1985) 85: 4397-4401).

An array of oligonucleotides probes comprising polynucleotide sequences or fragments thereof of the genes set forth in Table I can be constructed to conduct efficient screening of *e.g.*, genetic mutations. Such arrays are preferably high density arrays or grids. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability, see, for example, M. Chee et al., Science, 274, 610-613 (1996) and other references cited therein.

Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radio-immunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit comprising:

(a) a polynucleotide of the present invention, preferably the nucleotide sequence set forth in the Sequence Listing , or a fragment or an RNA transcript thereof;

(b) a nucleotide sequence complementary to that of (a);

(c) a polypeptide of the present invention, preferably the polypeptide set forth in the Sequence Listing or a fragment thereof; or

(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide set forth in the Sequence Listing .

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly diseases of the invention, amongst others.

The polynucleotide sequences of the present invention are valuable for chromosome localisation studies. The sequences set forth in the Sequence Listing are specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical

position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (co-inheritance of physically adjacent genes). Precise human chromosomal localisations for a genomic sequence (gene fragment etc.) can be determined using Radiation Hybrid (RH) Mapping (Walter, M. Spillett, D., Thomas, P., Weissenbach, J., and Goodfellow, P., (1994) A method for constructing radiation hybrid maps of whole genomes, Nature Genetics 7, 22-28). A number of RH panels are available from Research Genetics (Huntsville, AL, USA) e.g. the GeneBridge4 RH panel (Hum Mol Genet 1996 Mar;5(3):339-46 A radiation hybrid map of the human genome. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillett D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN). To determine the chromosomal location of a gene using this panel, 93 PCRs are performed using primers designed from the gene of interest on RH DNAs. Each of these DNAs contains random human genomic fragments maintained in a hamster background (human / hamster hybrid cell lines). These PCRs result in 93 scores indicating the presence or absence of the PCR product of the gene of interest. These scores are compared with scores created using PCR products from genomic sequences of known location. This comparison is conducted at http://www.genome.wi.mit.edu/.

The polynucleotide sequences of the present invention are also valuable tools for tissue expression studies. Such studies allow the determination of expression patterns of polynucleotides of the present invention which may give an indication as to the expression patterns of the encoded polypeptides in tissues, by detecting the mRNAs that encode them. The techniques used are well known in the art and include in situ hydridization techniques to clones arrayed on a grid, such as cDNA microarray hybridization (Schena *et al*, Science, 270, 467-470, 1995 and Shalon *et al*, Genome Res, 6, 639-645, 1996) and nucleotide amplification techniques such as PCR. A preferred method uses the TAQMAN (Trade mark) technology available from Perkin Elmer. Results from these studies can provide an indication of the normal function of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by an alternative form of the same gene (for example, one having an alteration in polypeptide coding potential or a regulatory mutation) can provide valuable insights into the role of the polypeptides of the present invention, or that of inappropriate expression thereof in disease. Such inappropriate expression may be of a temporal, spatial or simply

quantitative nature.

A further aspect of the present invention relates to antibodies. The polypeptides of the invention or their fragments, or cells expressing them, can be used as immunogens to produce antibodies that are immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against polypeptides of the present invention may also be employed to treat diseases of the invention, amongst others.

Polypeptides and polynucleotides of the present invention may also be used as vaccines. Accordingly, in a further aspect, the present invention relates to a method for inducing an immunological response in a mammal that comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said animal from disease, whether that disease is already established within the individual or not. An immunological response in a mammal may also be induced by a method comprises delivering a polypeptide of the present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases of the invention. One way of administering the vector is by accelerating it into the desired

cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid. For use a vaccine, a polypeptide or a nucleic acid vector will be normally provided as a vaccine formulation (composition). The formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intra-muscular, intravenous, or intra-dermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions that may contain anti-oxidants, buffers, bacteriostats and solutes that render the formulation instonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions that may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Polypeptides of the present invention have one or more biological functions that are of relevance in one or more disease states, in particular the diseases of the invention hereinbefore mentioned. It is therefore useful to identify compounds that stimulate or inhibit the function or level of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those that stimulate or inhibit the function or level of the polypeptide. Such methods identify agonists or antagonists that may be employed for therapeutic and prophylactic purposes for such diseases of the invention as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, collections of chemical compounds, and natural product mixtures. Such agonists or antagonists so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; a structural or functional mimetic thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)) or a small molecule. Such small molecules preferably have a molecular weight below 2,000 daltons, more preferably between 300 and 1,000 daltons, and most preferably between 400 and 700 daltons. It is preferred that these small molecules are organic molecules.

The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof, by means of a label directly or indirectly associated with the candidate compound.

Alternatively, the screening method may involve measuring or detecting (qualitatively or quantitatively) the competitive binding of a candidate compound to the polypeptide against a labeled competitor (*e.g.* agonist or antagonist). Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring an activity of the genes set forth in Table I in the mixture, and comparing activity of the mixture of the genes set forth in Table I to a control mixture which contains no candidate compound.

Polypeptides of the present invention may be employed in conventional low capacity screening methods and also in high-throughput screening (HTS) formats. Such HTS formats include not only the well-established use of 96- and, more recently, 384-well micotiter plates but also emerging methods such as the nanowell method described by Schullek et al, Anal Biochem., 246, 20-29, (1997).

Fusion proteins, such as those made from Fc portion and polypeptide of the genes set forth in Table I, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents that may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

A polypeptide of the present invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, ^{125}I), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell

supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide that compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

Examples of antagonists of polypeptides of the present invention include antibodies or, in some cases, oligonucleotides or proteins that are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, *e.g.*, a fragment of the ligands, substrates, receptors, enzymes, etc.; or a small molecule that bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Screening methods may also involve the use of transgenic technology and the genes set forth in Table I. The art of constructing transgenic animals is well established. For example, the genes set forth in Table I may be introduced through microinjection into the male pronucleus of fertilized oocytes, retroviral transfer into pre- or post-implantation embryos, or injection of genetically modified, such as by electroporation, embryonic stem cells into host blastocysts. Particularly useful transgenic animals are so-called "knock-in" animals in which an animal gene is replaced by the human equivalent within the genome of that animal. Knock-in transgenic animals are useful in the drug discovery process, for target validation, where the compound is specific for the human target. Other useful transgenic animals are so-called "knock-out" animals in which the expression of the animal ortholog of a polypeptide of the present invention and encoded by an endogenous DNA sequence in a cell is partially or completely annulled. The gene knock-out may be targeted to specific cells or tissues, may occur only in certain cells or tissues as a consequence of the limitations of the technology, or may occur in all, or substantially all, cells in the animal. Transgenic animal technology also offers a whole animal expression-cloning system in which introduced genes are expressed to give large amounts of polypeptides of the present invention

Screening kits for use in the above described methods form a further aspect of the present invention. Such screening kits comprise:

(a) a polypeptide of the present invention;

(b) a recombinant cell expressing a polypeptide of the present invention;

(c) a cell membrane expressing a polypeptide of the present invention; or

(d) an antibody to a polypeptide of the present invention;

which polypeptide is preferably that set forth in the Sequence Listing.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Glossary

The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an

Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Secreted protein activity or secreted polypeptide activity" or "biological activity of the secreted protein or secreted polypeptide" refers to the metabolic or physiologic function of said secreted protein including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said secreted protein.

"Secreted protein gene" refers to a polynucleotide comprising any of the attached nucleotide sequences or allelic variants thereof and/or their complements.

"Polynucleotide" generally refers to any polyribonucleotide (RNA) or polydeoxribonucleotide (DNA), which may be unmodified or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of

modifications may be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any polypeptide comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques that are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, Proteins - Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., Post-translational Protein Modifications: Perspectives and Prospects, 1-12, in Post-translational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", Meth Enzymol, 182, 626-646, 1990, and Rattan *et al.*, "Protein Synthesis: Post-

translational Modifications and Aging", Ann NY Acad Sci, 663, 48-62, 1992).

"Fragment" of a polypeptide sequence refers to a polypeptide sequence that is shorter than the reference sequence but that retains essentially the same biological function or activity as the reference polypeptide. "Fragment" of a polynucleotide sequence refers to a polynucleotide sequence that is shorter than the reference sequence set forth in the Sequence Listing.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains the essential properties thereof. A typical variant of a polynucleotide differs in nucleotide sequence from the reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from the reference polypeptide. Generally, alterations are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, insertions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. Typical conservative substitutions include Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe and Tyr. A variant of a polynucleotide or polypeptide may be naturally occurring such as an allele, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis. Also included as variants are polypeptides having one or more post-translational modifications, for instance glycosylation, phosphorylation, methylation, ADP ribosylation and the like. Embodiments include methylation of the N-terminal amino acid, phosphorylations of serines and threonines and modification of C-terminal glycines.

"Allele" refers to one of two or more alternative forms of a gene occurring at a given locus in the genome.

"Polymorphism" refers to a variation in nucleotide sequence (and encoded polypeptide sequence, if relevant) at a given position in the genome within a population.

"Single Nucleotide Polymorphism" (SNP) refers to the occurrence of nucleotide variability at a single nucleotide position in the genome, within a population. An SNP may occur within a gene or within intergenic regions of the genome. SNPs can be assayed using Allele Specific Amplification (ASA). For the process at least 3 primers are required. A

common primer is used in reverse complement to the polymorphism being assayed. This common primer can be between 50 and 1500 bps from the polymorphic base. The other two (or more) primers are identical to each other except that the final 3' base wobbles to match one of the two (or more) alleles that make up the polymorphism. Two (or more) PCR reactions are then conducted on sample DNA, each using the common primer and one of the Allele Specific Primers.

"Splice Variant" as used herein refers to cDNA molecules produced from RNA molecules initially transcribed from the same genomic DNA sequence but which have undergone alternative RNA splicing. Alternative RNA splicing occurs when a primary RNA transcript undergoes splicing, generally for the removal of introns, which results in the production of more than one mRNA molecule each of that may encode different amino acid sequences. The term splice variant also refers to the proteins encoded by the above cDNA molecules.

"Identity" reflects a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, determined by comparing the sequences. In general, identity refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of the two polynucleotide or two polypeptide sequences, respectively, over the length of the sequences being compared.

"% Identity" - For sequences where there is not an exact correspondence, a "% identity" may be determined. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or very similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length.

"Similarity" is a further, more sophisticated measure of the relationship between two polypeptide sequences. In general, "similarity" means a comparison between the amino acids of two polypeptide chains, on a residue by residue basis, taking into account not only exact correspondences between a between pairs of residues, one from each of the sequences being compared (as for identity) but also, where there is not an exact correspondence, whether, on an evolutionary basis, one residue is a likely substitute for the other. This likelihood has an associated "score" from which the "% similarity" of the two sequences can then be determined.

Methods for comparing the identity and similarity of two or more sequences are

well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J et al, Nucleic Acids Res, 12, 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity and the % similarity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (J Mol Biol, 147,195-197, 1981, Advances in Applied Mathematics, 2, 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences that are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences, finding a "maximum similarity", according to the algorithm of Neddleman and Wunsch (J Mol Biol, 48, 443-453, 1970). GAP is more suited to comparing sequences that are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3, for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S F et al, J Mol Biol, 215, 403-410, 1990, Altschul S F et al, Nucleic Acids Res., 25:389-3402, 1997, available from the National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA and accessible through the home page of the NCBI at www.ncbi.nlm.nih.gov) and FASTA (Pearson W R, Methods in Enzymology, 183, 63-99, 1990; Pearson W R and Lipman D J, Proc Nat Acad Sci USA, 85, 2444-2448,1988, available as part of the Wisconsin Sequence Analysis Package).

Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S and Henikoff J G, Proc. Nat. Acad Sci. USA, 89, 10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a reference polynucleotide or a polypeptide sequence, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value, as hereinbefore described.

"Identity Index" is a measure of sequence relatedness which may be used to compare a candidate sequence (polynucleotide or polypeptide) and a reference sequence.

Thus, for instance, a candidate polynucleotide sequence having, for example, an Identity Index of 0.95 compared to a reference polynucleotide sequence is identical to the reference sequence except that the candidate polynucleotide sequence may include on average up to five differences per each 100 nucleotides of the reference sequence. Such differences are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion. These differences may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between these terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polynucleotide sequence having an Identity Index of 0.95 compared to a reference polynucleotide sequence, an average of up to 5 in every 100 of the nucleotides of the in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

Similarly, for a polypeptide, a candidate polypeptide sequence having, for example, an Identity Index of 0.95 compared to a reference polypeptide sequence is identical to the reference sequence except that the polypeptide sequence may include an average of up to five differences per each 100 amino acids of the reference sequence. Such differences are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. These differences may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between these terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polypeptide sequence having an Identity Index of 0.95 compared to a reference polypeptide sequence, an average of up to 5 in every 100 of the amino acids in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

The relationship between the number of nucleotide or amino acid differences and the Identity Index may be expressed in the following equation:

$$n_a \leq x_a - (x_a \bullet I),$$

in which:

n_a is the number of nucleotide or amino acid differences,

x_a is the total number of nucleotides or amino acids in a sequence set forth in the Sequence Listing,

I is the Identity Index,

• is the symbol for the multiplication operator, and

in which any non-integer product of x_a and I is rounded down to the nearest integer prior to subtracting it from x_a.

"Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a reference sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the two sequences as hereinbefore defined. Falling within this generic term are the terms "ortholog", and "paralog". "Ortholog" refers to a polynucleotide or polypeptide that is the functional equivalent of the polynucleotide or polypeptide in another species. "Paralog" refers to a polynucleotideor polypeptide that within the same species which is functionally similar.

"Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. In one example, EP-A-0 464 533-A discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties [see, *e.g.*, EP-A 0232 262]. On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

Table I.

Gene Name	GSK Gene ID	Nucleic Acid SEQ ID NO's	Corresponding Protein SEQ ID NO's
sbg300828GLY	300828	SEQ ID NO:1 SEQ ID NO:2	SEQ ID NO:25 SEQ ID NO:26
sbg290600OLF	290600	SEQ ID NO:3	SEQ ID NO:27
sbg224366CALa	224366	SEQ ID NO:4 SEQ ID NO:5	SEQ ID NO:28 SEQ ID NO:29
sbg317645CRF	317645	SEQ ID NO:6	SEQ ID NO:30
sbg323398LYS	323398	SEQ ID NO:7	SEQ ID NO:31
sbg222729Cda	222729	SEQ ID NO:8 SEQ ID NO:9	SEQ ID NO:32 SEQ ID NO:33
sbg313227VDCCa	313227	SEQ ID NO:10 SEQ ID NO:11	SEQ ID NO:34 SEQ ID NO:35
sbg327427mia	327427	SEQ ID NO:12	SEQ ID NO:36
sbg318729proa	318729	SEQ ID NO:13 SEQ ID NO:14	SEQ ID NO:37 SEQ ID NO:38
sbg263419CARa	263419	SEQ ID NO:15 SEQ ID NO:16	SEQ ID NO:39 SEQ ID NO:40
sbg334109TES	334109	SEQ ID NO:17 SEQ ID NO:18	SEQ ID NO:41 SEQ ID NO:42
sbg323357SRCR	sbg323357	SEQ ID NO:19	SEQ ID NO:43
sbg294576LAPP	294576	SEQ ID NO:20	SEQ ID NO:44
sbg320795MMPa	320795	SEQ ID NO:21 SEQ ID NO:22	SEQ ID NO:45 SEQ ID NO:46
sbh312883.PLK	312883	SEQ ID NO:23	SEQ ID NO:47
sbg66804SPARCra	66804	SEQ ID NO:24	SEQ ID NO:48

Table II

Gene Name	Gene Family	Closest Polynuclotide by homology	Closest Polypeptide by homology	Cell Localization (by homology)
sbg300828-GLY	Proteoglycan	SC:DJ994D16 Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Human GROS1-L protein, gi:11127638, Kaul,S.C., Sugihara,T., Yoshida,A., Nomura,H. and Wadhwa,R. Oncogene 19 (32), 3576-3583 (2000)	Secreted
sbg290600-OLF	Olfactomedin-related protein	SC:BA292C23 Submitted by Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK	Rat neuronal olfactomedin-related ER localized protein precursor, GB:Q62609, Danielson,P.E., Forss-Petter,S., Battenberg,E.L., deLecea,L., Bloom,F.E., and Sutcliffe,J.G., 1994, J. Neurosci. Res. 38:468-478	Secreted
sbg224366-CALa	Cadherin	GB:AC006203 Submitted (18-DEC-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Human cadeherin 20, gi:10834607, Kools,P., Van Imschoot,G. and van Roy,F. Genomics 68 (3), 283-295 (2000)	Secreted
sbg317645-CRF	C1q-related factor (CRF)	GB:AC019017 Submitted (28-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA.	Human C1q-related factor, GI:5729785, Berube NG, Swanson XH, Bertram MJ, Kittle JD, Didenko V, Baskin DS, Smith JR and Pereira-Smith OM., 1999, Brain Res. Mol. Brain Res. 63:233-240.	Secreted
sbg323398-LYS	Lysozyme C precursor	GB:Z98304, Submitted (12-MAY-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK	Human Hydrolase protein-1, geneseqp: Y52597, Submitted by INCYTE PHARM INC, Publication number and date: WO200028045-A2, 18-MAY-00	Secreted
sbg222729-Cda	Leukocyte differentiation antigen	GB:AC012471 Submitted (28-OCT-1999) by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA	Mouse lymphocyte antigen 108 isoforms, gi:9887091, Submitted (21-MAR-2000) Department of Microbiology and Immunology, Vanderbilt University School of Medicine, 1161 21st Ave South / AA4206 Medical Center North, Nashville, TN 37232-2363, USA	Secreted
sbg313227-VDCCa	Voltage-dependent calcium channel	GB:AC005342 and GB:AC005343 Both were submitted (31-JUL-1998) by Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Mouse calcium channel alpha2delta, gi: 6753236, Klugbauer,N., Lacinova,L., Marais,E., Hobom,M. and Hofmann,F., J. Neurosci. 19, 648-691 (1999)	Membrane-bound
sbg327427-MIA	Melanoma inhibitory activity protein	SC:AL034428 Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK	Human melanoma derived growth regulatory protein precursor, gi:2498559 Blesch A, Bosserhoff AK, Apfel R, Behl C, Hessdoerfer B, Schmitt A, Jachimczak P, Lottspeich F, Buettner R, Bogdahn U, 1994, Cancer Res. 54:5695-5701.	Secreted

Table II (cont).

Gene Name	Gene Family	Closest Polynuclotide by homology	Closest Polypeptide by homology	Cell Localization (by homology)
sbg318729-proa	2-19 protein precursor	GB:AC022471 Submitted (04-FEB-2000) by Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA	Human 2-19 protein precursor gi:2135170 Bione S, Tamanini F, Maestrini E, Tribioli C, Poustka A, Torri G, Rivella S, Toniolo D. Transcriptional organization of a 450-kb region of the human X chromosome in Xq28. Proc Natl Acad Sci U S A 1993 Dec 1; 90(23): 10977-81	Secreted
sbg263419-CARa	Carboxypeptidase A1	GB:AC007938 Submitted (01-JUL-1999) by Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA.	Pig carboxypeptidase A1, gi:4336196, Submitted (02-JUL-1998) by LBBN, CNRS-UPRESA 6033, Faculte des Sciences et Techniques de St. Jerome, Universite d'Aix-Marseille, Av. Escadrille Normandie Niemen, Marseille 13397, France	Cytosolic
sbg334109-TES	Testatin precursor	GB:AL121894 Submitted (17-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Mouse testatin precursor (cystatin 9), gi:6753546 Tohonen V, Osterlund C, and Nordqvist K, 1998, Proc Natl Acad Sci USA 95:14208-13.	Secreted
sbg323357-SRCR	Scavenger receptor cysteine-rich (SRCR)	GB:AL161645 Submitted (17-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Bovine WC1 antigen, gi:26741, Wijngaard PL, Metzelaar MJ, MacHugh ND, Morrison WI, and Clevers HC, 1992, J. Immunol. 149:3273-3277.	Membrane-bound
sbg294576-LAPP	Lysosomal acid phosphatase precursor	JGI: CITB-E1_2568A17 Joint Genome Institute, Department of Energy, USA	Mouse lysosomal acid phosphatase precursor, gi:130728, Geier C, von Figura K, and Pohlmann R, 1991, Biol Chem Hoppe Seyler 372:301-4.	Secreted
sbg320795-MMPa	Matrix metalloproteinase	GB:AL158835 Submitted (05-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK	Xenopus laevis matrix metalloproteinase gene, gi:3211705, Yang,M., Murray,M.T. and Kurkinen,M., A novel matrix metalloproteinase gene (XMMP) encoding vitronectin-like motifs is transiently expressed in Xenopus laevis early embryo development. 1997 J. Biol. Chem. 272 (21), 13527-13533	Secreted
sbh312883.-PLK	Proteoglycan link protein (PLK)	GB:AC003967 Submitted (31-DEC-1997) by Human Genome Center, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA	Chicken cartilage link protein, gi:130309, Deak, F., Kiss,I., Sparks,K.J., Argraves,W.S., Hampikian,G. and Goetinck,P.F, Proc. Natl. Acad. Sci. U.S.A. 83 (11), 3766-3770 (1986)	Secreted
sbg66804-SPARCra	Sparc-related protein	GB:AL135747 Submitted by Genoscope – Centre National de Sequencage :BP 19191006 EVRY cedex, France	Mouse SPARC-related rpotein, gi:5305327 Submitted (05-Jun-1998) by GeneCraft, Treskowst. 10, Muenster 48163, Germany.	Membrane-bound

Table III

Gene Name	Uses	Associated Diseases
sbg300828-GLY	An embodiment of the invention is the use of sbg300828GLY, a proteoglycan, to control the sequence of ganglion cell differentiation and initial direction of axons and/or the differentiation of cells during development and maintenance of tissue organization. Proteoglycans are complex glycoconjugates containing a core protein to which a variable number of glycosaminoglycan chains (such as heparin sulfate, chondroitin sulfate, etc.) are covalently attached (Hassel J.R., Kimura J.H., and Hascall V.C., 1986, Annu. Rev. Biochem. 55:539-567). Interactions between negatively charged glycosaminoglycan chains and molecules such as growth factors are essential for differentiation of cells during development and maintenance of tissue organization (Prydz K, and Dalen KT, 2000, J Cell Sci 113:193-205). It has also been reported that in the developing retina a chondroitin sulfate proteoglycan appears to play an essential role in controlling the sequence of ganglion cell differentiation and initial direction of axons (Silver J, 1994, J Neurol 242:S22-4).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, and inflammation.
sbg290600-OLF	An embodiment of the invention is the use of sbg290600OLF, a glycoprotein, in chemoreception and the central nervous system. A close homologue of sbg290600OLF is olfactomedin. Olfactomedin is a glycoprotein, and reacts with proteins of olfactory cilia. It was originally discovered at the mucociliary surface of the amphibian olfactory neuroepithelium and subsequently found throughout the mammalian brain (Danielson,P.E., Forss-Petter,S., Battenberg,E.L., deLecea,L., Bloom,F.E., and Sutcliffe,J.G., 1994, J. Neurosci. Res. 38:468-478). Its noticeable deposition at the chemosensory surface of the olfactory neuroepithelium suggests a role for this protein in chemoreception (Snyder DA, Rivers AM, Yokoe H, Menco BP, and Anholt RR, 1991, Biochemistry 30:9143-53). The widespread occurrence of olfactomedin among mammalians also suggests its new functions in the central nervous system (Karavanich CA, and Anholt RR, 1998, Mol Biol Evol 15:718-26).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, and inflammation.
sbg224366-CALa	An embodiment of the invention is the use of sbg224366CALa, a secreted protein, in the identification of targets for new cancer therapies. A close homologue of sbg224366CALa is the mouse cadherin 7 precursor. The cadherins are calcium dependent cell adhesion proteins that preferentially interact with themselves in a homophilic mannerin connecting cells; cadherins may contribute to the sorting of heterogeneous cell types and is claimed to be involved in tumor progression. (Faulkner-Jones,B.E., Godhino,L.N.M., Pasquini,G.F., Reese,B.E. and Tan,S.-S. Cloning And Expression Of Mouse Cadherin-7, A Type-II Cadherin Isolated From the Developing Eye. Molecular and Cellular Neurosciences. Mol. Cell. Neurosci. (1999) In press).	Infections, cancers, autoimmune disorders, wound healing disorders, and hematopoietic disorders.
sbg317645-CRF	An embodiment of the invention is the use of sbg317645CRF in functions of the central nervous system, particularly the brain and motor functions. A close homologue of sbg224366CALa is C1q. C1q is a subunit of the C1 enzyme complex that activates the serum complement system. It has been shown that human CRF transcript is expressed at highest levels in the brain, particularly in the brainstem. Similarly, in mouse brain CRF transcripts are most abundant in areas of the nervous system involved in motor function (Berube NG, Swanson XH, Bertram MJ, Kittle JD, Didenko V, Baskin DS, Smith JR, and Pereira-Smith OM., 1999, Brain Res. Mol. Brain Res. 63:233-240).	Nervous system disorder.
sbg323398-LYS	An embodiment of the invention is the use of sbg323398LYS, a lysozyme, to inhance the activity of immunoagents in tissue and body fluids. Lysozymes are originally a bacteriolytic defensive agent and has been adapted to serve a digestive function (Qasba PK, Kumar S, 1997, Crit Rev Biochem Mol Biol 32:255-306). It has been suggested that lysozymes may serve as biomarkers of periodontal disease activity from inflammatory cell origin (Eley BM, and Cox SW, 1998, Br Dent J 184:323-8).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, and inflammation.

Table III (cont).

Gene Name	Uses	Associated Diseases
sbg222729-CDa	An embodiment of the invention is the use of sbg222729Cda, a secreted protein, in the diagnosis and treatment of cancer and autoimmune disorders. A close homologue of sbg222729Cda is leukocyte differentiation antigen CD84 isoform. CD84, a member of the immunoglobulin superfamily, shows high homology with several molecules belonging to the CD2 family of differentiation antigens, is proposed to be useful in the diagnosis and treatment of cancer and autoimmune disorders (Palou E, Pirotto F, Sole J, Freed JH, Peral B, Vilardell C, Vilella R, Vives J, Gaya A. Genomic characterization of CD84 reveals the existence of five isoforms differing in their cytoplasmic domains. Tissue Antigens 2000 Feb;55(2):118-27).	Cancer, autoimmune disorder, wound healing disorder, infections and hematopoietic disorders
sbg313227-VDCCa	An embodiment of the invention is the use of sbg313227-VDCCa in excitation-contraction coupling, and drug screening for obtaining agonists and antagonists. A close homologue of sbg313227-VDCCa is the calcium channel, voltage dependent, alpha2/delta subunit 3. The l-type calcium channel is composed of four subunits: alpha-1, alpha-2, beta and gamma. Alpha-2 and delta forms heterodimers that are disulfide-linked. Alpha2/delta-3 is expressed exclusively in the brain, e.g., in the hippocampus, cerebellum, and cortex, whereas alpha2/delta-2 is found in several tissues.	Cancer, Infections, autoimmune disorders, wound healing disorders and hematopoietic disorders
sbg327427-MIA	An embodiment of the invention is the use of sbg327427MIA, a growth regulating protein, as a future antitumor therapeutical agent. Close homologues of sbg327427MIA are melanoma inhibitory activity (MIA) proteins. MIA proteins have growth inhibition on melanoma cells in vitro as well as some other neuroectodermal tumors, including gliomas. (Blesch A, Bosserhoff AK, Apfel R, Behl C, Hessdoerfer B, Schmitt A, Jachimczak P, Lottspeich F, Buettner R, Bogdahn U, 1994, Cancer Res. 54:5695-5701).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, and inflammation.
sbg318729-PROa	An embodiment of the invention is the use of sbg318729PROa, a secreted protein, in the diagnosis and treatment of diseases of muscle and brain tissues. A close homologue of sbg318729PROa is the 2-19 protein precursor. The 2-19 protein maps to Xq28, is highly expressed in muscle and brain, and may be responsible for muscle or neurological disorders mapped to distal Xq28 (Bione S, Tamanini F, Maestrini E, Tribioli C, Poustka A, Torri G, Rivella S, Toniolo D. Transcriptional organization of a 450-kb region of the human X chromosome in Xq28. Proc Natl Acad Sci U S A 1993 Dec 1;90(23):10977-81).	Cancer, autoimmune disorders, infections, wound healing disorders and hematopoietic disorders
sbg263419-CARa	An embodiment of the invention is the use of sbg263419CARa in antibody-direct enzyme pro-drug therapy of viral infections. A close homologue of sbg263419CARa is human carboxypeptidase A1. Human carboxypeptidase A1 is useful in antibody-direct enzyme prodrug therapy of viral infections (MOORE JT, OHMSTEDE C and DEV IK, Molecular chimaera for use in enzyme gene therapy - is activated in a target cell to express a secretable enzyme which cleaves a prodrug outside the cell into a cytotoxic or cytostatic agent. Accession Number R97618. Publication Date: 30-MAY-96).	Infections, cancers, autoimmune disorders, wound healing disorders and hematopoietic disorders
sbg334109-TES	An embodiment of the invention is the use of sbg334109TES in natural tissue remodeling events such as bone resorption and embryo implantation and/or tumor formation and metastasis. A close homologue of sbg334109TES is testatin. Testatin is related to a group of cysteine protease inhibitors known as cystatins. Testatins and their target proteases can induce testis formation in foetal gonads, and may be associated with tumor formation and metastasis. In addition, it is suggested that they are also involved in natural tissue remodeling events such as bone resorption and embryo implantation (Tohonen V, Osterlund C, and Nordqvist K, 1998, Proc Natl Acad Sci USA 95:14208-13).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation, and infertility

Table III (cont).

Gene Name	Uses	Associated Diseases
sbg323357-SRCR	An embodiment of the invention is the use of sbg323357SRCR in receptor-mediated endocytosis of chemically modified lipoproteins and the pathogenesis of atherosclersis. Close homologues of sbg323357SRCR are scavenger receptors. Scavenger receptors are involved in receptor-mediated endocytosis of chemically modified lipoproteins, such as acetylated and oxidized LDL, and therefore have been implicated in the pathogenesis of atherosclersis (Adachi H, Tsujimoto M, Arai H, and Inoue K, 1997, J Biol Chem 272:31217-20). Especially, macrophage scavenger receptors have been implicated both in the deposition of lipoprotein cholesterol in artery walls during the formation of atherosclerotic plaques and in host defense against infections (Krieger M, 1992 Trends Biochem Sci 17:141-6).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, and inflammation
sbg294576-LAPP	An embodiment of the invention is the use of sbg294576LAPP in the diagnosis and treatment of prostatic cancer, osteolysis, Gaucher's disease of the spleen, and hairy cell leukemia. Close homologues of sbg294576LAPP are acid phosphatases. The acid phosphatases have been used as a marker for prostatic cancer, and have been linked with miscellaneous disorders, notably increased osteolysis, Gaucher's disease of spleen, and hairy cell leukemia (Moss DW, Raymond FD, and Wile DB; 1995; Crit Rev Clin Lab Sci 32:431-67).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation, increased osteolysis, and Gaucher's disease
sbg320795-MMPa	An embodiment of the invention is the use of sbg320795-MMPa, a secreted protein, in the treatment, prevention, and diagnosis of diabetic nephropathy, glomerulonephritis, fibrosis, liver cirrhosis, and metabolic bone diseases such as osteoporosis. A close homologue of sbg320795-MMPa is xenopus laevis matrix metalloproteinase. Xenopus laevis matrix metalloproteinase specifically activates pro-gelatinase a, which is involved in extracellular matrix turn-over on the surface of cells and is involved in the matrix remodeling of blood vessels (Yang,M., Murray,M.T. and Kurkinen,M., A novel matrix metalloproteinase gene (XMMP) encoding vitronectin-like motifs is transiently expressed in Xenopus laevis early embryo development. J. Biol. Chem. 272 (21), 13527-13533 (1997)).	Diabetic nephropathy, glomerulonephritis, fibrosis, liver cirrhosis and metabolic bone disease such as osteoporosis
sbh312883-PLK	An embodiment of the invention is the use of sbh312883-PLK to treat autoimmune diseases such as insulin dependent diabetes mellitus, multiple sclerosis, autoimmune thyroiditis, uveoretinitis, rheumatoid arthritis, and abnormal inflammatory immune responses. Close homologues of sbh312883-PLK are immunotherapeutic agents. Similar peptides have been used as antigen base immunotherapeutic agents in hosts afflicted with autoimmune diseases.	Hematopoietic disorders, wound healing disorders, viral and bacterial infection, cancer, and autoimmune diseases such as insulin dependent diabetes mellitus, multiple sclerosis, autoimmune thyroiditis, uveoretinitis, rheumatoid arthritis, and abnormal inflammatory immune responses
sbg66804-SPARCra	An embodiment of the invention is the use of sbg66804-SPARCra, a secreted protein, in remodeling, development, cell turnover, tissue repair, counter adhesion, and antiproliferation. A close homologue of sbg66804-SPARCra, is the mouse SPARC-related protein. SPARC (secreted protein, acidic and rich in cysteine) is a unique matricellular glycoprotein that is expressed by many different types of cells and is associated with development, remodeling, cell turnover, and tissue repair. Its principal functions in vitro are counter adhesion and antiproliferation, which proceed via different signaling pathways. SPARC has demonstrated activities in angiogenesis, cataractogenesis, and wound healing. SPARC has also been identified in tumors. The sequence of SPARC has been highly conserved among species.	Cataractogenesis, angiogenesis, wound healing, tumors.

Table IV. Quantitative, Tissue-specific mRNA expression detected using SybrMan

Quantitative, tissue-specific, mRNA expression patterns of the genes were measured using SYBR-Green Quantitative PCR (Applied Biosystems, Foster City, CA; see Schmittgen T.D. et al., Analytical Biochemistry 285:194-204, 2000) and human cDNAs prepared from various human tissues. Gene-specific PCR primers were designed using the first nucleic acid sequence listed in the Sequence List for each gene. Results are presented as the number of copies of each specific gene's mRNA detected in 1ng mRNA pool from each tissue. Two replicate mRNA measurements were made from each tissue RNA.

Gene Name	Tissue-Specific mRNA Expression (copies per ng mRNA; avg. ± range for 2 data points per tissue)									
	Brain	Heart	Lung	Liver	Kidney	Skeletal muscle	Intes tine	Spleen lymph	Placenta	Testis
sbg300828-GLY	2513 ±66	4268 ±154	4488 ±236	4229 ±250	4801 ±79	1801 ±29	2108 ±138	7431 ±152	15800 ±364	14682 ±1152
sbg290600-OLF	5164 ±119	234 ±19	266 ±41	88 ±13	378 ±43	187 ±115	177 ±23	159 ±31	239 ±27	292 ±4
sbg224366-CALa	636 ±34	13 ±4	6 ±1	-13 ±2	20 ±0	73 ±16	-1 ±1	3 ±1	-1 ±1	5 ±2
sbg323398-LYS	142 ±8	151 ±2	201 ±14	61 ±6	232 ±23	72 ±13	69 ±12	176 ±4	240 ±0	4015 ±251
sbg222729-CDa	12 ±1	50 ±2	304 ±2	50 ±8	100 ±6	145 ±4	166 ±4	2703 ±75	150 ±8	133 ±12
sbg313227-VDCCa	28 ±6	5 ±3	22 ±2	6 ±8	7 ±2	6 ±2	1 ±4	23 ±1	91 ±22	419 ±15
sbg263419-CARa	26 ±5	16 ±3	29 ±10	-2 ±6	42 ±4	143 ±3	3 ±1	112 ±11	177 ±10	8301 ±627
sbg323357-SRCR	131 ±8	78 ±7	131 ±20	57 ±5	193 ±18	107 ±3	59 ±1	178 ±3	197 ±50	181 ±47
sbg294576-LAPP	113 ±10	89 ±1	67 ±20	16 ±1	51 ±12	91 ±1	61 ±14	80 ±1	74 ±0	1618 ±117
sbg320795-MMPa	19 ±0	258 ±26	2886 ±114	219 ±7	367 ±27	168 ±19	4232 ±277	46644 ±1535	340 ±22	4160 ±205
sbg312883-PLK	364 ±4	3 ±3	3 ±0	96 ±11	8 ±0	4 ±2	22 ±2	-6 ±4	3 ±0	-5 ±7
sbg66804-SPARCra	296 ±53	24 ±0	4 ±1	457 ±21	7 ±0	68 ±3	9 ±1	439 ±11	128 ±1	1037 ±17

Table V. Additional diseases based on mRNA expression in specific tissues

Tissue Expression	Additional Diseases
Brain	Neurological and psychiatric diseases, including Alzheimers, parasupranuclear palsey, Huntington's disease, myotonic dystrophy, anorexia, depression, schizophrenia, headache, amnesias, anxiety disorders, sleep disorders, multiple sclerosis
Heart	Cardiovascular diseases, including congestive heart failure, dilated cardiomyopathy, cardiac arrhythmias, Hodgson's Disease, myocardial infarction, cardiac arrhythmias
Lung	Respiratory diseases, including asthma, Chronic Obstructive Pulmonary Disease, cystic fibrosis, acute bronchitis, adult respiratory distress syndrome
Liver	Dyslipidemia, hypercholesterolemia, hypertriglyceridemia, cirrhosis, hepatic encephalopathy, fatty hepatocirrhosis, viral and nonviral hepatitis, Type II Diabetes Mellitis, impaired glucose tolerance
Kidney	Renal diseases, including acute and chronic renal failure, acute tubular necrosis, cystinuria, Fanconi's Syndrome, glomerulonephritis, renal cell carcinoma, renovascular hypertension
Skeletal muscle	Eulenburg's Disease, hypoglycemia, obesity, tendinitis, periodic paralyses, malignant hyperthermia, paramyotonia congenita, myotonia congenita
Intestine	Gastrointestinal diseases, including Myotonia congenita, Ileus, Intestinal Obstruction, Tropical Sprue, Pseudomembranous Enterocolitis
Spleen/lymph	Lymphangiectasia, hypersplenism, angiomas, ankylosing spondylitis, Hodgkin's Disease, macroglobulinemia, malignant lymphomas, rheumatoid arthritis
Placenta	Choriocarcinoma, hydatidiform mole, placenta previa
Testis	Testicular cancer, male reproductive diseases, including low testosterone and male infertility
Pancreas	Diabetic ketoacidosis, Type 1 & 2 diabetes, obesity, impaired glucose tolerance

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

(a) an isolated polypeptide encoded by a polynucleotide comprising a sequence set forth in Table I;

(b) an isolated polypeptide comprising a polypeptide sequence set forth in Table I; and

(c) a polypeptide sequence of a gene set forth in Table I.

2. An isolated polynucleotide selected from the group consisting of:

(a) an isolated polynucleotide comprising a polynucleotide sequence set forth in Table I;

(b) an isolated polynucleotide of a gene set forth in Table I;

(c) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide set forth in Table I;

(d) an isolated polynucleotide encoding a polypeptide set forth in Table I;

(e) a polynucleotide which is an RNA equivalent of the polynucleotide of (a) to (d);

or a polynucleotide sequence complementary to said isolated polynucleotide.

3. An expression vector comprising a polynucleotide capable of producing a polypeptide of claim 1 when said expression vector is present in a compatible host cell.

4. A process for producing a recombinant host cell which comprises the step of introducing an expression vector comprising a polynucleotide capable of producing a polypeptide of claim 1 into a cell such that the host cell, under appropriate culture conditions, produces said polypeptide.

5. A recombinant host cell produced by the process of claim 4.

6. A membrane of a recombinant host cell of claim 5 expressing said polypeptide.

7. A process for producing a polypeptide which comprises culturing a host cell of claim 5 under conditions sufficient for the production of said polypeptide and recovering said polypeptide from the culture.

SEQUENCE LISTING

<110> SMITHKLINE BEECHAM CORPORATION
SMITHKLINE BEECHAM p.l.c.

<120> NOVEL COMPOUNDS

<130> GP50022

<140> TO BE ASSIGNED
<141> 2001-04-11

<150> 60/196,603
<151> 2000-04-13

<150> 60/199,417
<151> 2000-04-24

<160> 48

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2127
<212> DNA
<213> Homo sapiens

<400> 1
```
atggcggtac gcgcgttgaa gctgctgacc acactgctgg ctgtcgtggc cgctgcctcc      60
caagccgagg tcgagtccga ggcaggatgg ggcatggtga cgcctgatct gctcttcgcc     120
gaggggaccg cagcctacgc gcgcggggac tggcccgggg tggtcctgag catggaacgg     180
gcgctgcgct cccgggcagc cctccgcgcc cttcgcctgc gctgccgcac ccagtgtgcc     240
gccgacttcc cgtgggagct ggaccccgac tggtccccca gcccggccca ggcctcgggc     300
gccgccgccc tgcgcgacct gagcttcttc gggggccttc tgcgtcgcgc tgcctgcctg     360
cgccgctgcc tcgggccgcc ggccgcccac tcgctcagcg aagagatgga gctggagttc     420
cgcaagcgga gcccctacaa ctacctgcag gtcgcctact tcaagatcaa caagttggag     480
aaagctgttg ctgcagcaca caccttcttc gtgggcaatc ctgagcacat ggaaatgcag     540
cagaacctag actattacca aaccatgtct ggagtgaagg aggccgactt caaggatctt     600
gagactcaac cccatatgca agaatttcga ctgggagtgc gactctactc agaggaacag     660
ccacaggaag ctgtgcccca cctagaggcg gcgctgcaag aatactttgt ggcctatgag     720
gagtgccgtg ccctctgcga agggccctat gactacgatg gctacaacta ccttgagtac     780
```

aacgctgacc tcttccaggc catcacagat cattacatcc aggtcctcaa ctgtaagcag 840
aactgtgtca cggagcttgc ttcccaccca agtcgagaga agccctttga agacttcctc 900
ccatcgcatt ataattatct gcagtttgcc tactataaca agacaatctg ctattgtaat 960
cttccttgtc ttctgaaaat ctatagaaaa aagaagagtg ccaaggagta ccgacagcga 1020
agcctactgg aaaaagaact gcttttcttc gcttatgatg tttttggaat tccctttgtg 1080
gatccggatt catggactcc agaagaagtg attcccaaga gattgcaaga gaaacagaag 1140
tcagaacggg aaacagccgt acgcatctcc caggagattg ggaaccttat gaaggaaatc 1200
gagacccttg tggaagagaa gaccaaggag tcactggatg tgagcagact gacccgggaa 1260
ggtggccccc tgctgtatga aggcatcagt ctcaccatga actccaaact cctgaatggt 1320
tcccagcggg tggtgatgga cggcgtaatc tctgaccacg agtgtcagga gctgcagaga 1380
ctgaccaatg tggcagcaac ctcaggagat ggctaccggg gtcagacctc cccacatact 1440
cccaatgaaa agttctatgg tgtcactgtc ttcaaagccc tcaagctggg gcaagaaggc 1500
aaagttcctc tgcagagtgc ccacctgtac tacaacgtga cggagaaggt gcggcgcatc 1560
atggagtcct acttccgcct ggatacgccc ctctactttt cctactctca tctggtgtgc 1620
cgcactgcca tcgaagaggt ccaggcagag aggaaggatg atagtcatcc agtccacgtg 1680
gacaactgca tcctgaatgc cgagaccctc gtgtgtgtca aagagccccc agcctacacc 1740
ttccgcgact acagcgccat cctttaccta aatggggact tcgatggcgg aaacttttat 1800
ttcactgaac tggatgccaa gaccgtgacg gcagaggtgc agcctcagtg tggaagagcc 1860
gtgggattct cttcaggcac tgaaaaccca catggagtga aggctgtcac caggggggcag 1920
cgctgtgcca tcgccctgtg gttcaccctg gaccctcgac acagcgagcg ggacagggtg 1980
caggcagatg acctggtgaa gatgctcttc agcccagaag agatggacct ctcccaggag 2040
cagcccctgg atgcccagca gggtcccccc gaacctgcac aagagtctct ctcaggcagt 2100
gaatcgaagc ccaaggatga gctatga 2127

<210> 2
<211> 2211
<212> DNA
<213> Homo sapiens

<400> 2
atggcggtac gcgcgttgaa gctgctgacc acactgctgg ctgtcgtggc cgctgcctcc 60
caagccgagg tcgagtccga ggcaggatgg ggcatggtga cgcctgatct gctcttcgcc 120
gaggggaccg cagcctacgc gcgcggggac tggcccgggg tggtcctgag catggaacgg 180
gcgctgcgct cccgggcagc cctccgcgcc cttcgcctgc gctgccgcac ccagtgtgcc 240
gccgacttcc cgtgggagct ggaccccgac tggtccccca gcccggccca ggcctcgggc 300
gccgccgccc tgcgcgacct gagcttcttc gggggccttc tgcgtcgcgc tgcctgcctg 360
cgccgctgcc tcgggccgcc ggccgcccac tcgctcagcg aagagatgga gctggagttc 420
cgcaagcgga gcccctacaa ctacctgcag gtcgcctact tcaagatcaa caagttggag 480
aaagctgttg ctgcagcaca caccttcttc gtgggcaatc ctgagcacat ggaaatgcag 540
cagaacctag actattacca aaccatgtct ggagtgaagg aggccgactt caaggatctt 600
gagactcaac cccatatgca agaatttcga ctgggagtgc gactctactc agaggaacag 660
ccacaggaag ctgtgcccca cctagaggcg gcgctgcaag aatactttgt ggcctatgag 720

gagtgccgtg ccctctgcga agggccctat gactacgatg gctacaacta ccttgagtac 780
aacgctgacc tcttccaggc catcacagat cattacatcc aggtcctcaa ctgtaagcag 840
aactgtgtca cggagcttgc ttcccaccca agtcgagaga agccctttga agacttcctc 900
ccatcgcatt ataattatct gcagtttgcc tactataaca ttgggaatta tacacaggct 960
gttgaatgtg ccaagaccta tcttctcttc ttccccaatg acgaggtgat gaaccaaaat 1020
ttggcctatt atgcagctat gcttggagaa gaacacacca gatccatcgg cccccgtgag 1080
agtgccaagg agtaccgaca gcgaagccta ctggaaaaag aactgctttt cttcgcttat 1140
gatgtttttg gaattccctt tgtggatccg gattcatgga ctccagaaga agtgattccc 1200
aagagattgc aagagaaaca gaagtcagaa cgggaaacag ccgtacgcat ctcccaggag 1260
attgggaacc ttatgaagga aatcgagacc cttgtggaag agaagaccaa ggagtcactg 1320
gatgtgagca gactgacccg ggaaggtggc cccctgctgt atgaaggcat cagtctcacc 1380
atgaactcca aactcctgaa tggttcccag cgggtggtga tggacggcgt aatctctgac 1440
cacgagtgtc aggagctgca gagactgacc aatgtggcag caacctcagg agatggctac 1500
cggggtcaga cctccccaca tactcccaat gaaaagttct atggtgtcac tgtcttcaaa 1560
gccctcaagc tggggcaaga aggcaaagtt cctctgcaga gtgcccacct gtactacaac 1620
gtgacggaga aggtgcggcg catcatggag tcctacttcc gcctggatac gcccctctac 1680
ttttcctact ctcatctggt gtgccgcact gccatcgaag aggtccaggc agagaggaag 1740
gatgatagtc atccagtcca cgtggacaac tgcatcctga atgccgagac cctcgtgtgt 1800
gtcaaagagc ccccagccta caccttccgc gactacagcg ccatccttta cctaaatggg 1860
gacttcgatg gcggaaactt ttatttcact gaactggatg ccaagaccgt gacggcagag 1920
gtgcagcctc agtgtggaag agccgtggga ttctcttcag gcactgaaaa cccacatgga 1980
gtgaaggctg tcaccagggg gcagcgctgt gccatcgccc tgtggttcac cctggaccct 2040
cgacacagcg agcgggacag ggtgcaggca gatgacctgg tgaagatgct cttcagccca 2100
gaagagatgg acctctccca ggagcagccc ctggatgccc agcagggtcc ccccgaacct 2160
gcacaagagt ctctctcagg cagtgaatcg aagcccaagg atgagctatg a 2211

<210> 3
<211> 1437
<212> DNA
<213> Homo sapiens

<400> 3
atgagtcctc cactgctgaa gcttggcgct gtgcttagta ccatggcaat gatctcaaac 60
tggatgtccc aaactctccc atccttggtg ggactgaaca ccacgaggct gtcgactccg 120
gataccttaa ctcagattag tcctaaagaa gggtggcagg tgtacagctc agctcaggat 180
cctgatgggc ggtgcatttg cacagttgtt gctccagaac aaaacctgtg ttcccgggat 240
gccaaaagca ggcaacttcg ccaactactg gaaaaggttc agaacatgtc ccagtctatt 300
gaagtcttaa acttgagaac tcagagagat ttccaatatg ttttaaaaat ggaaacccaa 360
atgaaagggc tgaaggcaaa atttcggcag attgaagatg atcgaaagac acttatgacc 420
aagcattttc aggagttgaa agagaaaatg gacgagctcc tgcctttgat ccccgtgctg 480
gaacagtaca aaacagatgc taagttaatc acccagttca aggaggaaat aaggaatctg 540
tctgctgtcc tcactggtat tcaggaggaa attggtgcct atgactacga ggaactacac 600

caaagagtgc tgagcttgga aacaagactt cgtgactgca tgaaaaagct aacatgtggc 660
aaactgatga aaatcacagg cccagttaca gtcaagacat ctggaacccg atttggtgct 720
tggatgacag accctttagc atctgagaaa aacaacagag tctggtacat ggacagttat 780
actaacaata aaattgttcg tgaatacaaa tcaattgcag actttgtcag tggggctgaa 840
tcaaggacat acaaccttcc tttcaagtgg gcaggaacta accatgttgt ctacaatggc 900
tcactctatt ttaacaagta tcagagtaat atcatcatca aatacagctt tgatatgggg 960
agagtgcttg cccaacgaag cctggagtat gctggttttc ataatgttta cccctacaca 1020
tggggtggat tctctgacat cgacctaatg gctgatgaaa tcgggctgtg ggctgtgtat 1080
gcaactaacc agaatgcagg caatattgtc atcagccaac ttaaccaaga taccttggag 1140
gtgatgaaga gctggagcac tggctacccc aagagaagtg caggggaatc tttcatgatc 1200
tgtgggacac tgtatgtcac caactcccac ttaactggag ccaaggtgta ttattcctat 1260
tccaccaaaa cctccacata tgagtacaca gacattccct tccataacca atactttcac 1320
atatccatgc ttgactacaa tgcaagagat cgagctctct atgcctggaa caatggccac 1380
caggtgctgt tcaatgtcac ccttttccat atcatcaaga cagaggatga cacatag 1437

<210> 4
<211> 1770
<212> DNA
<213> Homo sapiens

<400> 4
atgtggactt ctggtagaat gagcaatgca aagaactggc ttggacttgg catgtccttg 60
tacttctggg ggctgatgga ccttacgacc accgttctct cggacacccc aacaccacaa 120
ggtgaattag aagcactcct gtcagacaag ccacagtcac atcagcggac caagaggagc 180
tgggtttgga accagttttt cgttctggaa gagtacactg ggaccgaccc tttgtatgtc 240
ggcaagcttc attcagatat ggacagggga gacggatcca tcaaatacat cctctcggga 300
gaaggtgctg gcatcgtgtt taccatcgac gacaccactg gagacatcca cgccattcag 360
aggctcgacc gagaggaaag agcccagtat actctaaggg ctcaagccct agacaggcgg 420
acgggcaggc caatggagcc cgagtcagag ttcatcatca aaattcaaga catcaatgac 480
aatgagccca agttcctgga cggaccttat gtggccactg tgccagaaat gtcccctgtg 540
ggtacctccg tcatccaagt gacagccaca gatgcagatg acccgaccta cggcaacagt 600
gccagggtgg tgtacagcat tcttcagggc cagccatatt tttctgtgga ctctaaaaca 660
ggtgtaatta ggacagcgct catgaacatg gacagagaag ccaaagaata ctacgaagtg 720
attatccaag ccaaggacat gggagggcag cttggaggat tagctgggac cacaacagtc 780
aacatcaccc tctcagatgt caatgataac ccaccccgct ttccccagaa acattaccag 840
atgagtgtgt tggaatcagc tccaattagc tccactgtcg ggagagtgtt tgccaaggac 900
ttggatgaag gcatcaatgc agagatgaaa tatactattg tggatggaga tggtgcagat 960
gcctttgaca ttagcacaga tcccaatttc caagttggta tcataactgt gaagaagccc 1020
ctgagttttg aaagcaagaa aagctacacc ttaaaggtgg agggagccaa tcctcaccta 1080
gagatgcgtt ttctgaactt gggcccattt caggacacaa caacagtgca catcagtgtg 1140
gaagacgtgg acgagccccc tgtgtttgaa cctggctttt actttgtgga ggtgcctgag 1200
gatgtggcga ttggaacaac catacagatc atttctgcca aggacccaga tgtgaccaac 1260

aactcaatca gatactccat tgatagaagc agtgaccctg gaagattttt ctatgttgac 1320
attacaacag gtgccctaat gacagcaaga cccctagacc gggaagaatt ttcttggcat 1380
aatatcactg tccttgctat ggaaatgaac aatccctccc aggttggaag tgttcctgtc 1440
acaatcaaag tcttagatgt gaatgacaat gctccagagt tccccagatt ctatgaagct 1500
tttgtctgtg agaacgccaa ggcaggacag ctgatccaga cagtgagtgc ggtggaccaa 1560
gatgacccac gcaatggtca gcatttctac tacagcttgg ctcctgaggc tgctaacaac 1620
cccaacttta ccataaggga caaccaaggt aatcaggtgg atggttggct atctgtgctt 1680
ttctacagca taggccagct actttgggtt actgtcttat gcaaacagtg tcaaaggcta 1740
cctgttccat accagcaggg aggatgttaa 1770

<210> 5
<211> 2406
<212> DNA
<213> Homo sapiens

<400> 5
atgtggactt ctggtagaat gagcaatgca aagaactggc ttggacttgg catgtccttg 60
tacttctggg ggctgatgga ccttacgacc accgttctct cggacacccc aacaccacaa 120
ggtgaattag aagcactcct gtcagacaag ccacagtcac atcagcggac caagaggagc 180
tgggtttgga accagttttt cgttctggaa gagtacactg ggaccgaccc tttgtatgtc 240
ggcaagcttc attcagatat ggacagggga gacggatcca tcaaatacat cctctcggga 300
gaaggtgctg gcatcgtgtt taccatcgac gacaccactg gagacatcca cgccattcag 360
aggctcgacc gagaggaaag agcccagtat actctaaggg ctcaagccct agacaggcgg 420
acgggcaggc caatggagcc cgagtcagag ttcatcatca aaattcaaga catcaatgac 480
aatgagccca agttcctgga cggaccttat gtggccactg tgccagaaat gtcccctgtg 540
ggtacctccg tcatccaagt gacagccaca gatgcagatg acccgaccta cggcaacagt 600
gccagggtgg tgtacagcat tcttcagggc cagccatatt tttctgtgga ctctaaaaca 660
ggtgtaatta ggacagcgct catgaacatg gacagagaag ccaaagaata ctacgaagtg 720
attatccaag ccaaggacat gggagggcag cttggaggat tagctgggac cacaacagtg 780
aacatcaccc tctcagatgt caatgataac ccacccccgct ttccccagaa acattaccag 840
atgagtgtgt tggaatcagc tccaattagc tccactgtcg ggagagtgtt tgccaaggac 900
ttggatgaag gcatcaatgc agagatgaaa tatactattg tggatggaga tggtgcagat 960
gcctttgaca ttagcacaga tcccaatttc caagttggta tcataactgt gaagaagccc 1020
ctgagttttg aaagcaagaa aagctacacc ttaaaggtgg agggagccaa tcctcaccta 1080
gagatgcgtt ttctgaactt gggcccattt caggacacaa caacagtgca catcagtgtg 1140
gaagacgtgg acgagccccc tgtgtttgaa cctggctttt actttgtgga ggtgcctgag 1200
gatgtggcga ttggaacaac catacagatc atttctgcca aggacccaga tgtgaccaac 1260
aactcaatca gatactccat tgatagaagc agtgaccctg gaagattttt ctatgttgac 1320
attacaacag gtgccctaat gacagcaaga cccctagacc gggaagaatt ttcttggcat 1380
aatatcactg tccttgctat ggaaatgaac aatccctccc aggttggaag tgttcctgtc 1440
acaatcaaag tcttagatgt gaatgacaat gctccagagt tccccagatt ctatgaagct 1500
tttgtctgtg agaacgccaa ggcaggacag ctgatccaga cagtgagtgc ggtggaccaa 1560

gatgacccac gcaatggtca gcatttctac tacagcttgg ctcctgaggc tgctaacaac 1620
cccaacttta ccataaggga caaccaagat aacacagcac ggattctaac caggaggtct 1680
ggtttccggc agcaggagca gagtgtcttt cacctgccta tcctgatagc agatagcggg 1740
cagcccgtgc tgagcagcac aggcacactg accatccaag tgtgcagctg tgatgacgac 1800
ggccacgtca tgtcctgcag cccagaggcc tacatgctcc cagtcagttt gagccggggc 1860
gccctcattg ccatcctcgc ctgcatcttt gtcctcttag tgctggtgtt gctcattttg 1920
tccatgaggc ggcaccggaa acaaccatac atcatcgacg acgaggaaaa catccacgag 1980
aacatcgtcc gctacgacga cgagggcggc ggcgaggagg acaccgaggc cttcgacatc 2040
gcggccatgt ggaacccccg ggaggcgcag gcgggggccg cccccaagac gcggcaggac 2100
atgctgcccg agatcgagag cctctcccgc tacgtgcctc agacgtgcgc agtgaacagc 2160
actgtccaca gctacgtgct ggccaagctc tacgaggccg acatggacct gtgggcaccg 2220
cccttcgact ccctccagac gtatatgttc gagggggacg gctctgtggc ggggtcgctg 2280
agctccctgc agtcggccac gtcggactcg gaacagagct tcgacttcct gacggactgg 2340
gggccccgct tccggaagct ggccgagctc tacggggcgt cggagggacc cgcgccgctg 2400
tggtga 2406

<210> 6
<211> 864
<212> DNA
<213> Homo sapiens

<400> 6
atggcactgg ggctgctgat cgcggtgcct ctgctgctgc aggcggcgcc ccccggagcg 60
gctcactacg agatgctggg cacctgccgc atgatctgtg acccatacag cgtcgctccc 120
gcagggggac ccgcgggcgc caaggctcca ccgccgggac ccagtaccgc tgccctggaa 180
gttatgcagg acctcagcgc caaccccccg cctccgttta tccagggacc aaagggtgat 240
ccggggcgac caggcaagcc agggcctcgg ggtcctcctg gagagccagg gcctcctggg 300
cccaggggtc ccccgggaga gaaaggagac tcggggaggc cagggctacc cggactgcag 360
ttgacaacca gcgcggccgg tggcgttgga gtggtgagtg gcggaaccgg gggcggtggc 420
gacacggagg gagaagtgac cagtgcgctg agcgccgcct tcagcggtcc caagatcgcc 480
ttctacgtgg gactcaagag cccccacgaa ggctacgagg tgctcaagtt cgacgacgtg 540
gtcaccaatc ttggcaatca ctacgacccc actacaggca agttcagctg ccaggtgcgg 600
ggcatctact tcttcacgta ccacattctc atgcgtggcg gcgacggaac cagcatgtgg 660
gcggatctct gcaagaacgg gcaggtgcga gccagcgcca tagcccagga cgcggaccag 720
aattacgact acgccagcaa cagcgtggta ctgcacctgg attcaggcga tgaagtctac 780
gtgaagctgg acggcgggaa ggctcacggc ggcaacaata acaagtacag cacgttctcg 840
ggcttcctcc tgtatccgga ttag 864

<210> 7
<211> 480
<212> DNA
<213> Homo sapiens

<400> 7

```
atgaaggcct ggggcactgt ggtagtgacc ttggccacgc tgatggttgt cactgtggat      60
gccaagatct atgaacgctg cgagctggcg gcaagactgg agagagcagg gctgaacggc     120
tacaagggct acggcgttgg agactggctg tgcatggctc attatgagag tggctttgac     180
accgccttcg tggaccacaa tcctgatggc agcagtgaat atggcatttt ccaactgaat     240
tctgcctggt ggtgtgacaa tggcattaca cccaccaaga acctctgcca catggattgt     300
catgacctgc tcaatcgcca tattctggat gacatcaggt gtgccaagca gattgtgtcc     360
tcacagaatg ggctttctgc ctggacttct tggaggctac actgttctgg ccatgattta     420
tctgaatggc tcaaggggtg tgatatgcat gtgaaaattg atccaaaaat tcatccatga     480
```

<210> 8
<211> 663
<212> DNA
<213> Homo sapiens

<400> 8

```
atggtcagga acatttttaa aacctttcct tctgtgttta cagggaatgt agtttcacaa      60
agcagcttaa ccccattgat ggtgaacggg attctggggg agtcagtaac tcttcccctg     120
gagtttcctg caggagagaa ggtcaacttc atcacttggc ttttcaatga aacatctctt     180
gccttcatag taccccatga aaccaaaagt ccagaaatcc acgtgactaa tccgaaacag     240
ggaaagcgac tgaacttcac ccagtcctac tccctgcaac tcagcaacct gaagatggaa     300
gacacaggct cttacagagc ccagatatcc acaaagacct ctgcaaagct gtccagttac     360
actctgagga tattaagaca actgaggaac atacaagtta ccaatcacag tcagctattt     420
cagaatatga cctgtgagct ccatctgact tgctctgtgg aggatgcaga tgacaatgtc     480
tcattcagat gggaggcctt gggaaacaca ctttcaagtc agccaaacct cactgtctcc     540
tgggacccca ggatttccag tgaacaggac tacacctgca tagcagagaa tgctgtcagt     600
aatttatcct tctctgtctc tgcccagaag ctttgcgaag gtaacagcct gcctcaggtc     660
tga                                                                   663
```

<210> 9
<211> 1041
<212> DNA
<213> Homo sapiens

<400> 9

```
atgactgcct caaggtctca agcaccagtc ttcaccgcgg aaagcatgtt gtggctgttc      60
caatcgctcc tgtttgtctt ctgctttggc ccagggaatg tagtttcaca aagcagctta     120
accccattga tggtgaacgg gattctgggg gagtcagtaa ctcttcccct ggagtttcct     180
gcaggagaga aggtcaactt catcacttgg cttttcaatg aaacatctct tgccttcata     240
gtaccccatg aaaccaaaag tccagaaatc cacgtgacta atccgaaaca gggaaagcga     300
ctgaacttca cccagtccta ctccctgcaa ctcagcaacc tgaagatgga agacacaggc     360
```

tcttacagag cccagatatc cacaaagacc tctgcaaagc tgtccagtta cactctgagg 420
atattaagac aactgaggaa catacaagtt accaatcaca gtcagctatt tcagaatatg 480
acctgtgagc tccatctgac ttgctctgtg gaggatgcag atgacaatgt ctcattcaga 540
tgggaggcct tgggaaacac actttcaagt cagccaaacc tcactgtctc ctgggacccc 600
aggatttcca gtgaacagga ctacacctgc atagcagaga atgctgtcag taatttatcc 660
ttctctgtct ctgcccagaa gctttgcgaa gatgttaaaa ttcaatatac agataccaaa 720
atgattctgt ttatggtttc tgggatatgc atagtcttcg gtttcatcat actgctgtta 780
cttgttttga ggaaaagaag agattcccta tctttgtcta ctcagcgaac acagggcccc 840
gagtccgcaa ggaacctaga gtatgtttca gtgtctccaa cgaacaacac tgtgtatgct 900
tcagtcactc attcaaacag ggaaacagaa atctggacac ctagagaaaa tgatactatc 960
acaatttact ccacaattaa tcattccaaa gagagtaaac ccactttttc cagggcaact 1020
gcccttgaca atgtcgtgta a 1041

<210> 10
<211> 3228
<212> DNA
<213> Homo sapiens

<400> 10
atgggcacgg cttatctctg ctgtcctcaa gtgctcctcc tcctctgcct gccccggaga 60
gtgaagctat gggctgacac cttcggcggg gacctgtata acactgtgac caaatactca 120
ggctctctct tgctgcagaa gaagtacaag gatgtggagt ccagtctgaa gatcgaggag 180
gtggatggct tggagctggt gaggaagttc tcagaggaca tggagaacat gctgcggagg 240
aaagtcgagg cggtccagaa tctggtggaa gctgccgagg aggccgacct gaaccacgaa 300
ttcaatgaat ccctggtgtt cgactattac aactcggtcc tgatcaacga gagggacgag 360
aagggcaact tcgtggagct gggcgccgag ttcctcctgg agtccaatgc tcacttcagc 420
aacctgccgg tgaacacctc catcagcagc gtgcagctgc ccaccaacgt gtacaacaaa 480
gacccagata ttttaaatgg agtctacatg tctgaagcct tgaatgctgt cttcgtggag 540
aacttccaga gagacccaac gttgacctgg caatattttg gcagtgcaac tggattcttc 600
aggatctatc caggtataaa atggacacct gatgagaatg gagtcattac ttttgactgc 660
cgaaaccgcg gctggtacat tcaagctgct acttctccca aggacatagt gattttggtg 720
gacgtgagcg gcagtatgaa ggggctgagg atgactattg ccaagcacac catcaccacc 780
atcttggaca ccctgggggа gaatgacttc attaatatca tagcgtacaa tgactacgtc 840
cattacatcg agccttgttt taaagggatc ctcgtccagg cggaccgaga caatcgagag 900
catttcaaac tgctggtgga ggagttgatg gtcaaaggtg tgggggtcgt ggaccaagcc 960
ctgagagaag ccttccagat cctgaagcag ttccaagagg ccaagcaagg aagcctctgc 1020
aaccaggcca tcatgctcat cagcgacggc gccgtggagg actacgagcc ggtgtttgag 1080
aagtataact ggccagactg taaggtccga gttttcactt acctcattgg gagagaagtg 1140
tcttttgctg accgcatgaa gtggattgca tgcaacaaca aaggctacta cacgcagatc 1200
tcaacgctgg cggacaccca ggagaacgtg atggaatacc tgcacgtgct cagccgcccc 1260
atggtcatca accacgacca cgacatcatc tggacagagg cctacatgga cagcaagctc 1320
ctcagctcgc aggctcagag cctgacactg ctcaccactg tggccatgcc agtcttcagc 1380

aagaagaacg aaacgcgatc ccatggcatt ctcctgggtg tggtgggctc agatgtggcc 1440
ctgagagagc tgatgaagct ggcgccccgg tacaagcttg gagtgcacgg atacgccttt 1500
ctgaacacca acaatggcta catcctctcc catcccgacc tccggcccct gtacagagag 1560
gggaagaaac taaaacccaa acctaactac aacagtgtgg atctctccga agtggagtgg 1620
gaagaccagg ctgaatctct gagaacagcc atgatcaata gggaaacagg tactctctcg 1680
atggatgtga aggttccgat ggataaaggg aagcgagttc ttttcctgac caatgactac 1740
ttcttcacgg acatcagcga caccccttc agtttggggg tggtgctgtc ccggggccac 1800
ggagaataca tccttctggg gaacacgtct gtggaagaag gcctgcatga cttgcttcac 1860
ccagacctgg ccctggccgg tgactggatc tactgcatca cagatattga cccagaccac 1920
cggaagctca gccagctaga ggccatgatc cgcttcctca ccaggaagga cccagacctg 1980
gagtgtgacg aggagctggt ccgggaggtg ctgtttgacg cggtggtgac agcccccatg 2040
gaagcctact ggacagcgct ggccctcaac atgtccgagg agtctgaaca cgtggtggac 2100
atggccttcc tgggcacccg ggctggcctc ctgagaagca gcttgttcgt gggctccgag 2160
aaggtctccg acaggaagtt cctgacacct gaggacgagg ccagcgtgtt caccctggac 2220
cgcttcccgc tgtggtaccg ccaggcctca gagcatcctg ctggcagctt cgtcttcaac 2280
ctccgctggg cagaaggacc agaaagtgcg ggtgaaccca tggtggtgac ggcaagcaca 2340
gctgtggcgg tgaccgtgga caagaggaca gccattgctg cagccgcggg cgtccaaatg 2400
aagctggaat tcctccagcg caaattctgg gcggcaacgc ggcagtgcag cactgtggat 2460
gggccgtgca cacagagctg cgaggacagt gatctggact gcttcgtcat cgacaacaac 2520
gggttcattc tgatctccaa gaggtcccga gagacgggaa gatttctggg ggaggtggat 2580
ggtgctgtcc tgacccagct gctcagcatg gggtgttca gccaagtgac tatgtatgac 2640
tatcaggcca tgtgcaaacc ctcgagtcac caccacagtg cagcccagcc cctggtcagc 2700
ccaatttctg ccttcttgac ggcgaccagg tggctgctgc aggagctggt gctgttcctg 2760
ctggagtgga gtgtctgggg ctcctggtac gacagagggg ccgaggccca caaacacaag 2820
aagcaggacc cgctgcagcc ctgcgacacg gagtaccccg tgttcgtgta ccagccggcc 2880
atccgggagg ccaacgggat cgtggagtgc gggccctgcc agaaggtatt tgtggtgcag 2940
cagattccca acagtaacct cctcctcctg gtgacagacc ccaccttctg cagaatgggc 3000
tccggtcctg agatattgac cttaacagtg gcttctgcac ataatgcctc tgtcaaatgt 3060
gaccggatgc gctcccagaa gctccgccgg cgaccagact cctgccacgc cttccatcca 3120
gaggagaatg cccaggactg cggcggcgcc tcggacacct cagcctcgcc gcccctactc 3180
ctgctgcctg tgtgtgcctg gggctactg ccccaactcc tgcggtga 3228

<210> 11
<211> 3345
<212> DNA
<213> Homo sapiens

<400> 11
atgcctgcaa ctcccaactt cctcgcaaac cccagctcca gcagccgctg gattcccctc 60
cagccaatgc ccgtggcctg ggcctttgtg cagaagacct cggccctcct gtggctgctg 120
cttctaggca cctccctgtc ccctgcgtgg ggacaggcca agattcctct ggaaacagtg 180
aagctatggg ctgacacctt cggcggggac ctgtataaca ctgtgaccaa atactcaggc 240

```
tctctcttgc tgcagaagaa gtacaaggat gtggagtcca gtctgaagat cgaggaggtg     300
gatggcttgg agctggtgag gaagttctca gaggacatgg agaacatgct gcggaggaaa     360
gtcgaggcgg tccagaatct ggtggaagct gccgaggagg ccgacctgaa ccacgaattc     420
aatgaatccc tggtgttcga ctattacaac tcggtcctga tcaacgagag ggacgagaag     480
ggcaacttcg tggagctggg cgccgagttc ctcctggagt ccaatgctca cttcagcaac     540
ctgccggtga acacctccat cagcagcgtg cagctgccca ccaacgtgta caacaaagac     600
ccagatattt taaatggagt ctacatgtct gaagccttga atgctgtctt cgtggagaac     660
ttccagagag acccaacgtt gacctggcaa tattttggca gtgcaactgg attcttcagg     720
atctatccag gtataaaatg gacacctgat gagaatggag tcattacttt tgactgccga     780
aaccgcggct ggtacattca agctgctact tctcccaagg acatagtgat tttggtggac     840
gtgagcggca gtatgaaggg gctgaggatg actattgcca agcacaccat caccaccatc     900
ttggacaccc tgggggagaa tgacttcatt aatatcatag cgtacaatga ctacgtccat     960
tacatcgagc cttgttttaa agggatcctc gtccaggcgg accgagacaa tcgagagcat    1020
ttcaaactgc tggtggagga gttgatggtc aaaggtgtgg gggtcgtgga ccaagccctg    1080
agagaagcct tccagatcct gaagcagttc caagaggcca agcaaggaag cctctgcaac    1140
caggccatca tgctcatcag cgacggcgcc gtggaggact acgagccggt gtttgagaag    1200
tataactggc cagactgtaa ggtccgagtt ttcacttacc tcattgggag agaagtgtct    1260
tttgctgacc gcatgaagtg gattgcatgc aacaacaaag gctactacac gcagatctca    1320
acgctggcgg acacccagga gaacgtgatg gaatacctgc acgtgctcag ccgccccatg    1380
gtcatcaacc acgaccacga catcatctgg acagaggcct acatggacag caagctcctc    1440
agctcgcagg ctcagagcct gacactgctc accactgtgg ccatgccagt cttcagcaag    1500
aagaacgaaa cgcgatccca tggcattctc ctgggtgtgg tgggctcaga tgtggccctg    1560
agagagctga tgaagctggc gccccggtac aagcttggag tgcacggata cgcctttctg    1620
aacaccaaca atggctacat cctctcccat cccgacctcc ggcccctgta cagagagggg    1680
aagaaactaa aacccaaacc taactacaac agtgtggatc tctccgaagt ggagtgggaa    1740
gaccaggctg aatctctgag aacagccatg atcaataggg aaacaggtac tctctcgatg    1800
gatgtgaagg ttccgatgga taaagggaag cgagttcttt tcctgaccaa tgactacttc    1860
ttcacggaca tcagcgacac cccttcagt ttgggggtgg tgctgtcccg ggccacgga     1920
gaatacatcc ttctggggaa cacgtctgtg gaagaaggcc tgcatgactt gcttcaccca    1980
gacctggccc tggccggtga ctggatctac tgcatcacag atattgaccc agaccaccgg    2040
aagctcagcc agctagaggc catgatccgc ttcctcacca ggaaggaccc agacctggag    2100
tgtgacgagg agctggtccg ggaggtgctg tttgacgcgg tggtgacagc cccccatggaa   2160
gcctactgga cagcgctggc cctcaacatg tccgaggagt ctgaacacgt ggtggacatg    2220
gccttcctgg gcacccgggc tggcctcctg agaagcagct tgttcgtggg ctccgagaag    2280
gtctccgaca ggaagttcct gacacctgag gacgaggcca gcgtgttcac cctggaccgc    2340
ttcccgctgt ggtaccgcca ggcctcagag catcctgctg gcagcttcgt cttcaacctc    2400
cgctgggcag aaggaccaga aagtgcgggt gaacccatgg tggtgacggc aagcacagct    2460
gtggcggtga ccgtggacaa gaggacagcc attgctgcag ccgcgggcgt ccaaatgaag    2520
ctggaattcc tccagcgcaa attctgggcg gcaacgcggc agtgcagcac tgtggatggg    2580
ccgtgcacac agagctgcga ggacagtgat ctggactgct tcgtcatcga caacaacggg    2640
ttcattctga tctccaagag gtcccgagag acgggaagat ttctggggga ggtggatggt    2700
gctgtcctga cccagctgct cagcatgggg gtgttcagcc aagtgactat gtatgactat    2760
```

caggccatgt gcaaaccctc gagtcaccac cacagtgcag cccagcccct ggtcagccca 2820
atttctgcct tcttgacggc gaccaggtgg ctgctgcagg agctggtgct gttcctgctg 2880
gagtggagtg tctggggctc ctggtacgac agaggggccg aggcccacaa acacaagaag 2940
caggacccgc tgcagccctg cgacacggag taccccgtgt tcgtgtacca gccggccatc 3000
cgggaggcca acgggatcgt ggagtgcggg ccctgccaga aggtatttgt ggtgcagcag 3060
attcccaaca gtaacctcct cctcctggtg acagacccca ccttctgcag aatgggctcc 3120
ggtcctgaga tattgacctt aacagtggct tctgcacata atgcctctgt caaatgtgac 3180
cggatgcgct cccagaagct ccgccggcga ccagactcct gccacgcctt ccatccagag 3240
gagaatgccc aggactgcgg cggcgcctcg gacacctcag cctcgccgcc cctactcctg 3300
ctgcctgtgt gtgcctgggg gctactgccc caactcctgc ggtga 3345

<210> 12
<211> 387
<212> DNA
<213> Homo sapiens

<400> 12
atggcaagaa tattgttact tttcctcccg ggtcttgtgg ctgtatgtgc tgtgcatgga 60
atatttatgg accgtctagc ttccaagaag ctctgtgcag atgatgagtg tgtctatact 120
atttctctgg ctagtgctca agaagattat aatgccccgg actgtagatt cattaacgtt 180
aaaaaagggc agcagatcta tgtgtactca aagctggtaa aagaaaatgg agctggagaa 240
ttttgggctg gcagtgttta tggtgatggc caggacgaga tgggagtcgt gggttatttc 300
cccaggaact tggtcaagga acagcgtgtg taccaggaag ctaccaagga agttcccacc 360
acggatattg acttcttctg cgagtaa 387

<210> 13
<211> 648
<212> DNA
<213> Homo sapiens

<400> 13
atgggtctca cctggatcct agtcaccatc ctcctaggtg gtcctggtgt tggccttcct 60
cgaattcagc agttcttcac cagcccagag aactcagtga ctgcagaacc aagggccagg 120
aagtacaaat gcggcctgcc ccagccttgt cctgaagagc acctgagctt tcgaatagtc 180
agcggggctg ccaatgtcat cgggcccaag atctgcctcg aggacaagat gctcatgagc 240
agcgtcaaag acaatgtggg ccgtggcctg aacatcgccc tggtgaatgg ggtcagtggt 300
gagctcctag aagccagagc ctttgacatg tgggctggag atgtcaatga tctcttgaag 360
ttcatccggc cactgcatga aggtaccctg gtgtttgtgg cttcctatga tgatccagct 420
accaagatga atgaagagac caggaagctt ttttctgagc tgggcagcag gaatgccaag 480
gatctagcct tccgtgacag ctgggtgttt gtgggagcca aaggtgtgca gaacaagagc 540
ccctttgagc agcatatgaa gaacagtaag cacaccaaca agtatgaggg ctggccagag 600
gccctggaga tggaaggctg tatccctcga aggagcatag cgggctag 648

<210> 14
<211> 693
<212> DNA
<213> Homo sapiens

<400> 14
atgaggttgg caggcccccct gcgcatcgtg gccctaatca tcattatggg tctcacctgg 60
atcctagtca ccatcctcct aggtggtcct ggtgttggcc ttcctcgaat tcagcagttc 120
ttcaccagcc cagagaactc agtgactgca gaaccaaggg ccaggaagta caaatgcggc 180
ctgccccagc cttgtcctga agagcacctg agctttcgaa tagtcagcgg ggctgccaat 240
gtcatcgggc ccaagatctg cctcgaggac aagatgctca tgagcagcgt caaagacaat 300
gtgggccgtg gcctgaacat cgccctggtg aatggggtca gtggtgagct cctagaagcc 360
agagcctttg acatgtgggc tggagatgtc aatgatctct tgaagttcat ccggccactg 420
catgaaggta ccctggtgtt tgtggcttcc tatgatgatc cagctaccaa gatgaatgaa 480
gagaccagga agcttttttc tgagctgggc agcaggaatg ccaaggatct agccttccgt 540
gacagctggg tgtttgtggg agccaaaggt gtgcagaaca agagcccctt tgagcagcat 600
atgaagaaca gtaagcacac caacaagtat gagggctggc cagaggccct ggagatggaa 660
ggctgtatcc ctcgaaggag catagcgggc tag 693

<210> 15
<211> 1311
<212> DNA
<213> Homo sapiens

<400> 15
atgcagggca cccctggagg cgggacgcgc cctgggccat cccccgtgga caggcggaca 60
ctcctggtct tcagctttat cctggcagca gctttgggcc aaatgaattt cacaggggac 120
caggttcttc gagtcctggc caaagatgag aagcagcttt cacttctcgg ggatctggag 180
ggcctgaaac cccagaaggt ggacttctgg cgtggcccag ccaggcccag cctccctgtg 240
gatatgagag ttcctttctc tgaactgaaa gacatcaaag cttatctgga gtctcatgga 300
cttgcttaca gcatcatgat aaaggacatc caggtgctgc tggatgagga aagacaggcc 360
atggcgaaat cccgccggct ggagcgcagc accaacagct tcagttactc atcataccac 420
accctggagg agatatatag ctggattgac aactttgtaa tggagcattc cgatattgtc 480
tcaaaaattc agattggcaa cagctttgaa aaccagtcca ttcttgtcct gaagttcagc 540
actggaggtt ctcggcaccc agccatctgg attgacactg gaattcactc ccgggagtgg 600
atcacccatg ccaccggcat ctggactgcc aataagattg tcagtgatta tggcaaagac 660
cgtgtcctga cagacatact gaatgccatg gacatcttca tagagctcgt cacaaaccct 720
gatgggtttg cttttaccca cagcatgaac cgcttatggc ggaagaacaa gtccatcaga 780
cctggaatct tctgcatcgg cgtggatctc aacaggaact ggaagtcggg ttttggagga 840
aatggttcta acagcaaccc ctgctcagaa acttatcacg ggccctcccc tcagtcggag 900
ccggaggtgg ctgccatagt gaacttcatc acagcccatg gcaacttcaa ggctctgatc 960

tccatccaca gctactctca gatgcttatg taccettacg gccgattgct ggagcccgtt 1020
tcaaatcaga gggagttgta cgatcttgcc aaggatgcgg tggaggcctt gtataaggtc 1080
catgggatcg agtacatttt tggcagcatc agcaccaccc tctatgtggc cagtgggatc 1140
accgtcgact gggcctatga cagtggcatc aagtacgcct tcagctttga gctccgggac 1200
actgggcagt atggcttcct gctgccggcc acacagatca tccccacggc ccaggagacg 1260
tggatggcgc ttcggaccat catggagcac accctgaatc acccctacta g 1311

<210> 16
<211> 1260
<212> DNA
<213> Homo sapiens

<400> 16
atgcggacac tcctggtctt cagctttatc ctggcagcag ctttgggcca aatgaatttc 60
acaggggacc aggttcttcg agtcctggcc aaagatgaga agcagctttc acttctcggg 120
gatctggagg gcctgaaacc ccagaaggtg gacttctggc gtggcccagc caggcccagc 180
ctccctgtgg atatgagagt tcctttctct gaactgaaag acatcaaagc ttatctggag 240
tctcatggac ttgcttacag catcatgata aaggacatcc aggtgctgct ggatgaggaa 300
agacaggcca tggcgaaatc ccgccggctg gagcgcagca ccaacagctt cagttactca 360
tcataccaca ccctggagga gatatatagc tggattgaca actttgtaat ggagcattcc 420
gatattgtct caaaaattca gattggcaac agctttgaaa accagtccat tcttgtcctg 480
aagttcagca ctggaggttc tcggcaccca gccatctgga ttgacactgg aattcactcc 540
cgggagtgga tcacccatgc caccggcatc tggactgcca ataagattgt cagtgattat 600
ggcaaagacc gtgtcctgac agacatactg aatgccatgg acatcttcat agagctcgtc 660
acaaaccctg atgggtttgc ttttacccac agcatgaacc gcttatggcg gaagaacaag 720
tccatcagac ctggaatctt ctgcatcggc gtggatctca acaggaactg gaagtcgggt 780
tttggaggaa atggttctaa cagcaacccc tgctcagaaa cttatcacgg gccctcccct 840
cagtcggagc cggaggtggc tgccatagtg aacttcatca cagcccatgg caacttcaag 900
gctctgatct ccatccacag ctactctcag atgcttatgt acccttacgg ccgattgctg 960
gagcccgttt caaatcagag ggagttgtac gatcttgcca aggatgcggt ggaggccttg 1020
tataaggtcc atgggatcga gtacattttt ggcagcatca gcaccaccct ctatgtggcc 1080
agtgggatca ccgtcgactg ggcctatgac agtggcatca agtacgcctt cagctttgag 1140
ctccgggaca ctgggcagta tggcttcctg ctgccggcca cacagatcat ccccacggcc 1200
caggagacgt ggatggcgct tcggaccatc atggagcaca ccctgaatca cccctactag 1260

<210> 17
<211> 360
<212> DNA
<213> Homo sapiens

<400> 17
atgtggagtc tgccgccgag cagggctctg tcctgtgcgc cactgctgct tctcttcagc 60

ttccagttcc tggttaccta tgcttggcgt ttccaagagg aagaggagtg gaatgaccaa 120
aaacaaattg ctgtttatct ccctcccacc ctggagtttg ccgtgtacac attcaacaag 180
cagagcaagg actggtatgc ctacaagctg gtgcctgtcc tggcttcctg gaaggagcag 240
gttgatgagc acatcctttt ctgcactagt gtccagcaca ggctgctgag tgatgggcag 300
gggtggcagc gtgtggggca gggcttaacc aggactcctg gttcaccatt tgtagtctaa 360

<210> 18
<211> 447
<212> DNA
<213> Homo sapiens

<400> 18
atgtcgagtc cgcagaggag gaaggctatg ccctgggcac tgtcactgct tctcatgggc 60
ttccagctcc tggtgactta tgcctggtgt tctgaagagg aaatgggtgg taataataaa 120
atagtccagg atcctatgtt cctcgccaca gtggagtttg ccttgaacac tttcaacgtg 180
cagagcaagg aggagcatgc ctacaggctg ttgcgcgtcc tgagttcatg gagggaggat 240
agcatggaca gaaagatggt gttctccatg aatctgcaac tgcgccaaac cgtatgtagg 300
aaatttgaag atgacattga caactgccct tttcaagaaa gcctggagct gaacaacact 360
ttcacctgct tcttcaccat cagcaccagg ccctggatga ctcagttcag cctcctgaac 420
aagacctgct tggagggatt ccactga 447

<210> 19
<211> 2697
<212> DNA
<213> Homo sapiens

<400> 19
atgagggcag ctctctggac cctgggactc gggccccttc ttctgaatct ctgggcagtc 60
cccattggtg gaccaggtgc tctgaggctg gcgtacagac acagcacgtg cgacggagtg 120
gtgttggtcc gacaccacgg ggcatgggga tacgtgtgca accaggagtg gacgctggca 180
gaggcctctg tcgtgtgcag gcagctgggc tgcggccctg ccgtgggcgc ccccaagtat 240
gtcccgctgc ctggagagat ggcccagccc tggcttcaca acgtgtcctg ccggggcaac 300
gagtcctccc tctgggagtg cagccttggc tcatggtgcc agagcccgtg cccccacgca 360
tgggtggtgg tcgcgctgtg ctccaacggc actttccggg agctccggct ggtgaagggc 420
cgcagtccct gcgcgggact ccccgagatc agaaacgtga atggggtgga ccgcctctgt 480
gtcctgcatg tggaggaggc catggtgttc tgccgggagc tggggtgcgg ccctgtgctc 540
caggcccccc gccgggacgt gggcgtcgtc aggaagtacc tggcctgcag ggggtaccgag 600
cccaccatcc gcagctgcag actggacaac aacttccgca gcggctgcga cctgcggctg 660
gacgcagagg tggtctgctc aggacacacc gaggcccgac tggtgggcgg cgagcacccc 720
tgcgccgggc gcctggaggt gacctggggc accgtctgtg atgcggccct ggacctggcc 780
acggcccacg tggtgtgccg ggagctgcag tgtggggcgg tcgtgtccac gcccgagggc 840
gcccgcttcg gccggggctc ggggccggtg tggacggagg ccttccgctg tgcgggcaac 900

gagtcgctgc tgttccactg cccacggggg cgtgggagcc agtgtgggca tggtcacgac 960
gcggggctca ggtgctcaga gttcaggatg gtcaacggca gcagcagctg tgagggccgc 1020
gtggagttcc aggtgcaggg gtcctgggca cccctctgtg ccacccactg ggacatagca 1080
gatgccaccg tcctctgcca ccagctcaac tgtggcaacg cggtggccgc acctggagga 1140
ggccattttg gggacgggga cgctgccatc tggcctgatg cctttcactg tgaggggaca 1200
gagtcctact tgtggaattg cccagtaagc accctggggg ccccggcctg tgccccggga 1260
aacacagcct ccgcggtctg ctcaggtctg gcccacgccc tgcgactgag ggaaggacag 1320
agccgctgtg acggccgcgt ggaggtctcc ctggatggcg tgtggggccg cgtcctggac 1380
gatgcctggg acctgcgcgg cgcgggcgtg gtgtgccggc aactcgggtg cagaggggcc 1440
cagcaagcct atgacgcacc tgcccccagc cgcggatccg tccaggtggc gctgagccgc 1500
gtgcgctgtc tgggcaccga aacccgcctg actcagtgca acgtgtccgc gaccctgcag 1560
gagcccgcgg ggacctcgcg ggacgccggc gtggtgtgct ccggtgaggt cggaaccgcg 1620
tcccccatgg cccgtcgcca cgggatcccg ggcgccctga ctctgtctct ccacagggag 1680
cctcagggtg cggctggccg cggggccggg gcgctgcacg gggcgcgtg gggcaccgtg 1740
tgtgacgatg cctgggacct gcgggacgcg cacgtggtct gcaggcagct gggctgtggc 1800
cgcgccctga gcgccctggg ggccgcacac ttcggagccg gggcagggcg catctggctg 1860
gacgagctgg gctgccaggg ccacgagtct gcgctgtggc agtgcccgtc ggcgggctgg 1920
gggcggcacg actggaggca caaggaggac gccggcgtct tctgctcaga gtcggtggct 1980
ctgaggctgc gaggtgggac ctgctgctgt gctgggtggc tggacgtgtt ctacaatggg 2040
acctggggcg ccatgtgcag caatgccctg aaggacctct ccttgtccat catctgcaag 2100
cagctggggt gtggggtgtg gggagtgggg ctggctggag aacaggccct tcccctcgcg 2160
ggcaccggga ccgcctgggt ggacaacatc gagtgccgca ggctgcccaa ctccactctg 2220
tggcaatgcc cttcccaccc atggcacccg cactcttgcg accttcgaga gcaggtctgg 2280
attacctgtg cagtgaccgc agccccttt gcagaggagg gcgcactgcg cgtgcgcggg 2340
ggcgaggacc gctgctccgg gcgcgtggag ctctggcacg cgggctcctg gggcaccgtg 2400
tgcgacgatg gctgggacct ggcggacgcg gaggtcgtgt gccgccagct gggctgtggt 2460
cgggccgtcg ccgccctggg ggccgccgcc tttggccctg gctccgggcc cgtgtggctg 2520
gacgaggtgg ggtgccgggg cagcgaggcg tccctgtggg gctgccctgc ggagcggtgg 2580
ggacgcggag accgcgcgca cgaggaggac gcgggcgtgc gctgctgggg tgagtggggg 2640
gcggtgggaa gtcggtcatg gggccggcag agggcgctgg gatggagtca gtcttga 2697

<210> 20
<211> 1281
<212> DNA
<213> Homo sapiens

<400> 20
atggccggcc tggggttttg gggccaccct gctggacctc tcctgctgct gctgctgctg 60
gtgctgccac cccgggccct gccagaagga cccctggtgt tcgtggctct ggtattccgc 120
catggcgacc gggccccgct ggcctcctac cccatggacc cacacaagga ggtggcctcc 180
accctgtggc cacgaggcct gggccagctg accacggagg gggtccgcca gcagctggag 240
ctgggccgct tcctgaggag ccgctacgag gccttcctga gtccggagta ccggcgggag 300

gaggtgtaca tccgcagcac ggactttgac cgcacgctgg agagtgccca ggccaacctt 360
gccgggctgt ttcccgaggc tgctccaggg agccccgagg cccgctggag gccgatcccg 420
gtgcacacgg tgcccgtggc tgaggataag ctgctgaggt tccccatgcg cagctgtccc 480
cgataccacg agctgctgcg ggaggccacc gaggccgccg agtaccagga ggccctggag 540
ggctggacgg gcttcctgag tcgcctggag aacttcacgg gactgtcgct ggttggagag 600
ccactgcgca gggcatggaa ggttctggac accctcatgt gccagcaagc ccacggtctt 660
ccactaccag cctgggcctc cccagatgtc ctgcggactc ttgcccagat ctcggctttg 720
gatattggag cccacgtggg cccaccccgg gcagcagaga aggcccagct gacagggggg 780
atcctgctga atgctatcct tgcaaacttc tcccgggtcc agcgcctggg gctgcccctc 840
aagatggtca tgtactcagc tcatgacagc accctgctgg ccctccaggg ggccctgggc 900
ctctatgatg gacacacccc gccatatgct gcctgcctcg gctttgagtt ccggaagcac 960
ctgggggaatc ccgccaaaga tggagggaat gtcaccgtct ccctcttcta ccgcaatgac 1020
tccgcccacc tgcccctgcc tctcagcctc cccgggtgcc cggccccctg tccactaggc 1080
cgcttctacc agctgactgc cccggcccgg cctcccgccc atggggtctc ctgccatggc 1140
ccctatgagg ctgccatccc cccagctcca gtggtgcccc tgctggccgg agctgtagct 1200
gtgctggtgg cactcagctt ggggctgggc ctgctggcct ggagaccagg gtgcctgcgg 1260
gccttggggg gccccgtgtg a 1281

<210> 21
<211> 1428
<212> DNA
<213> Homo sapiens

<400> 21
atgctcgccg cctccatctt ccgtccgaca ctgctgctct gctggctggc tgctccctgg 60
cccacccagc ccgagagtct cttccacagc cgggaccgct cggacctgga gccgtcccca 120
ctgcgccagg ccaagcccat tgccgacctc cacgctgctc agcggttcct gtccagatac 180
ggctggtcag ggtgtgggc ggcctggggg cccagtcccg aggggccgcc ggagaccccc 240
aagggcgccg ccctggccga ggcggtgcgc aggttccagc gggcgaacgc gctgccggcc 300
agcggggagc tggacgcggc caccctagcg gccatgaacc ggccgcgctg cggggtcccg 360
gacatgcgcc caccgccccc ctccgccccg ccttcgcccc cgggcccgcc ccccagagcc 420
cgctccaggc gctccccgcg ggcgccgctg tccttgtccc ggcggggttg gcagccccgg 480
ggctaccccg acggcggagc tgcccaggcc ttctccaaga ggacgctgag ctggcggctg 540
ctgggcgagg ccctgagcag ccaactgtcc gtggccgacc agcggcgcat tgtggcgctg 600
gccttcagga tgtggagcga ggtgacgccg ctggacttcc gcgaggacct ggccgccccc 660
ggggccgcgg tcgacatcaa gctgggcttt gggagaggct cctgtgaggg atcatttgat 720
actgcgtttg actggattcg caaagagaga aaccaatatg gagaggtgat ggtgagattt 780
agcacatatt tcttccgtaa cagctggtac tggctttatg aaaatcgaaa caataggaca 840
cgctatgggg accctatcca aatcctcact ggctggcctg gaatcccaac acacaacata 900
gatgcctttg ttcacatctg gacatggaaa agagatgaac gttatttttt tcaaggaaat 960
caatactgga gatatgacag tgacaaggat caggccctca cagaagatga acaaggaaaa 1020
agctatccca aattgatttc agaaggattt cctggcatcc caagtcccct agacacggcg 1080

ttttatgacc gaagacagaa gttaatttac ttcttcaagg agtcccttgt atttgcattt 1140
gatgtcaaca gaaatcgagt acttaattct tatccaaaga ggattactga agtttttcca 1200
gcagtaatac cacaaaatca tcctttcaga aatatagatt ccgcttatta ctcctatgca 1260
tacaactcca ttttcttttt caaaggcaat gcatactgga aggtagttaa tgacaaggac 1320
aaacaacaga attcctggct tcctgctaat ggcttatttc caaaaaagtt tatttcagag 1380
aagtggtttg atgtttgtga cgtccatatc tccacactga acatgtaa 1428

<210> 22
<211> 1590
<212> DNA
<213> Homo sapiens

<400> 22
atgctcgccg cctccatctt ccgtccgaca ctgctgctct gctggctggc tgctccctgg 60
cccacccagc ccgagagtct cttccacagc cgggaccgct cggacctgga gccgtcccca 120
ctgcgccagg ccaagcccat tgccgacctc cacgctgctc agcggttcct gtccagatac 180
ggctggtcag ggtgtgggc ggcctggggg cccagtcccg aggggccgcc ggagaccccc 240
aagggcgccg ccctggccga ggcggtgcgc aggttccagc gggcgaacgc gctgccggcc 300
agcggggagc tggacgcggc caccctagcg gccatgaacc ggccgcgctg cgggccccgg 360
ggctaccccg acggcggagc tgcccaggcc ttctccaaga ggacgctgag ctggcggctg 420
ctgggcgagg ccctgagcag ccaactgtcc gtggccgacc agcggcgcat tgtggcgctg 480
gccttcagga tgtggagcga ggtgacgccg ctggacttcc gcgaggacct ggccgccccc 540
ggggccgcgg tcgacatcaa gctgggcttt gggagaggcc ggcacctggg ctgtccgcgg 600
gccttcgatg ggagcgggca ggagtttgca cacgcctggc gcctaggtga cattcacttt 660
gacgacgacg agcacttcac acctcccacc agtgacacgg gcatcagcct tctcaaggtg 720
gccgtccatg aaattggcca tgtcctgggc ttgcctcaca cctacaggac gggatccata 780
atgcaaccaa attacattcc ccaggagcct gcctttgagt tggactggtc agacaggaaa 840
gcaattcaaa agctgtatgg ctcctgtgag ggatcatttg atactgcgtt tgactggatt 900
cgcaaagaga gaaaccaata tggagaggtg atggtgagat ttagcacata tttcttccgt 960
aacagctggt actggcttta tgaaaatcga aacaatagga cacgctatgg ggaccctatc 1020
caaatcctca ctggctggcc tggaatccca acacacaaca tagatgcctt tgttcacatc 1080
tggacatgga aaagagatga acgttatttt tttcaaggaa atcaatactg gagatatgac 1140
agtgacaagg atcaggccct cacagaagat gaacaaggaa aaagctatcc caaattgatt 1200
tcagaaggat ttcctggcat cccaagtccc ctagacacgg cgttttatga ccgaagacag 1260
aagttaattt acttcttcaa ggagtccctt gtatttgcat ttgatgtcaa cagaaatcga 1320
gtacttaatt cttatccaaa gaggattact gaagtttttc cagcagtaat accacaaaat 1380
catcctttca gaaatataga ttccgcttat tactcctatg catacaactc cattttcttt 1440
ttcaaaggca atgcatactg gaaggtagtt aatgacaagg acaaacaaca gaattcctgg 1500
cttcctgcta atggcttatt tccaaaaaag tttatttcag agaagtggtt tgatgtttgt 1560
gacgtccata tctccacact gaacatgtaa 1590

<210> 23

<211> 1209
<212> DNA
<213> Homo sapiens

<400> 23

```
atggtgtgcg ctcgggcggc cctcggtccc ggcgcgctct gggccgcggc ctggggcgtc      60
ctgctgctca cagcccctgc gggggcgcag cgtggccgga agaaggtcgt gcacgtgctg     120
gagggtgagt cgggctcggt agtggtacag acagcgcctg ggcaggtggt aagccaccgt     180
ggtggcacca tcgtcttgcc ctgccgctac cactatgagg cagccgccca cggtcacgac     240
ggcgtccggc tcaagtggac aaaggtggtg gacccgctgg ccttcaccga cgtcttcgtg     300
gcactaggcc cccagcaccg ggcattcggc agctaccgtg ggcgggctga gctgcagggc     360
gacgggcctg ggatgcctc cctggtcctc cgcaacgtca cgctgcaaga ctacgggcgc     420
tatgagtgcg aagtcaccaa tgagctggaa gatgacgctg gcatggtcaa gctggacctg     480
gaaggcgtgg tctttcccta ccaccccgt ggaggccgat acaagctgac cttcgcggag      540
gcgcagcgcg cgtgcgccga gcaggacggc atcctggcat ctgcagaaca gctgcacgcg     600
gcctggcgcg acggcctgga ctggtgcaac gcgggctggt tgcgcgacgg ctcagtgcaa     660
taccccgtga accggccccg ggagccctgc ggcggcctgg gggggaccgg gagtgcaggg     720
ggcggcggtg atgccaacgg gggcctgcgc aactacgggt atcgccataa cgccgaggaa     780
cgctacgacg ccttctgctt cacgtccaac ctgccggggc gcgtgttctt cctgaagccg     840
ctgcgacctg tacccttctc cggagctgcg cgcgcgtgtg ctgcgcgtgg cgcggccgtg     900
gccaaggtgg ggcagctgtt cgccgcgtgg aagctgcagc tgctagaccg ctgcaccgcg     960
ggttggctgg ccgatggcag tgcgcgctac cccatcgtga acccgcgagc gcgctgcgga    1020
ggccgcaggc ctggtgtgcg cagcctcggc ttcccggacg ccacccgacg gctcttcggc    1080
gtctactgct accgcgctcc aggagcaccg gacccggcac ctggcggctg gggctggggc    1140
tgggcgggcg gcggcggctg ggcagggggc gcgcgcgatc ctgctgcctg gacccctctg    1200
cacgtctag                                                             1209
```

<210> 24
<211> 1326
<212> DNA
<213> Homo sapiens

<400> 24

```
atgctgcccg cgcgctgcgc ccgcctgctc acgccccact tgctgctggt gttggtgcag      60
ctgtcccctg ctcgcggcca ccgcaccaca ggccccaggt ttctaataag tgaccgtgac     120
ccacagtgca acctccactg ctccaggact caacccaaac ccatctgtgc ctctgatggc     180
aggtcctacg agtccatgtg tgagtaccag cgagccaagt gccgagaccc gaccctgggc     240
gtggtgcatc gaggtagatg caaagatgct ggccagagca agtgtcgcct ggagcgggct     300
caagccctgg agcaagccaa gaagcctcag gaagctgtgt ttgtcccaga gtgtggcgag     360
gatggctcct ttacccaggt gcagtgccat acttacactg ggtactgctg gtgtgtcacc     420
ccggatggga agcccatcag tggctcttct gtgcagaata aaactcctgt atgttcaggt     480
tcagtcaccg acaagccctt gagccagggt aactcaggaa ggaaagatga cgggtctaag     540
```

```
ccgacaccca cgatggagac ccagccggtg ttcgatggag atgaaatcac agccccaact      600
ctatggatta aacacttggt gatcaaggac tccaaactga acaacaccaa cataagaaat      660
tcagagaaag tctattcgtg tgaccaggag aggcagagtg ccctggaaga ggcccagcag      720
aatccccgtg agggtattgt catccctgaa tgtgcccctg ggggactcta taagccagtg      780
caatgccacc agtccactgg ctactgctgg tgtgtgctgg tggacacagg gcgcccgctg      840
cctgggacct ccacacgcta cgtgatgccc agttgtgaga gcgacgccag ggccaagact      900
acagaggcgg atgacccctt caaggacagg gagctaccag gctgtccaga agggaagaaa      960
atggagttta tcaccagcct actggatgct ctcaccactg acatggttca ggccattaac     1020
tcagcagcgc ccactggagg tgggaggttc tcagagccag accccagcca caccctggag     1080
gagcgggtag tgcactggta tttcagccag ctggacagca atagcagcaa cgacattaac     1140
aagcgggaga tgaagccctt caagcgctac gtgaagaaga aagccaagcc caagaaatgt     1200
gcccggcgtt tcaccgacta ctgtgacctg aacaaagaca aggtcatttc actgcctgag     1260
ctgaagggct gcctgggtgt tagcaaagaa ggtggtagcc ttggcagttt cccccaggca     1320
aaatga                                                                1326
```

<210> 25
<211> 708
<212> PRT
<213> Homo sapiens

<400> 25

```
Met Ala Val Arg Ala Leu Lys Leu Leu Thr Thr Leu Leu Ala Val Val
1               5                   10                  15
Ala Ala Ala Ser Gln Ala Glu Val Glu Ser Glu Ala Gly Trp Gly Met
            20                  25                  30
Val Thr Pro Asp Leu Leu Phe Ala Glu Gly Thr Ala Ala Tyr Ala Arg
        35                  40                  45
Gly Asp Trp Pro Gly Val Val Leu Ser Met Glu Arg Ala Leu Arg Ser
    50                  55                  60
Arg Ala Ala Leu Arg Ala Leu Arg Leu Arg Cys Arg Thr Gln Cys Ala
65                  70                  75                  80
Ala Asp Phe Pro Trp Glu Leu Asp Pro Asp Trp Ser Pro Ser Pro Ala
                85                  90                  95
Gln Ala Ser Gly Ala Ala Ala Leu Arg Asp Leu Ser Phe Phe Gly Gly
            100                 105                 110
Leu Leu Arg Arg Ala Ala Cys Leu Arg Arg Cys Leu Gly Pro Pro Ala
        115                 120                 125
Ala His Ser Leu Ser Glu Glu Met Glu Leu Glu Phe Arg Lys Arg Ser
    130                 135                 140
Pro Tyr Asn Tyr Leu Gln Val Ala Tyr Phe Lys Ile Asn Lys Leu Glu
145                 150                 155                 160
Lys Ala Val Ala Ala Ala His Thr Phe Phe Val Gly Asn Pro Glu His
```

165 170 175
Met Glu Met Gln Gln Asn Leu Asp Tyr Tyr Gln Thr Met Ser Gly Val
180 185 190
Lys Glu Ala Asp Phe Lys Asp Leu Glu Thr Gln Pro His Met Gln Glu
195 200 205
Phe Arg Leu Gly Val Arg Leu Tyr Ser Glu Glu Gln Pro Gln Glu Ala
210 215 220
Val Pro His Leu Glu Ala Ala Leu Gln Glu Tyr Phe Val Ala Tyr Glu
225 230 235 240
Glu Cys Arg Ala Leu Cys Glu Gly Pro Tyr Asp Tyr Asp Gly Tyr Asn
245 250 255
Tyr Leu Glu Tyr Asn Ala Asp Leu Phe Gln Ala Ile Thr Asp His Tyr
260 265 270
Ile Gln Val Leu Asn Cys Lys Gln Asn Cys Val Thr Glu Leu Ala Ser
275 280 285
His Pro Ser Arg Glu Lys Pro Phe Glu Asp Phe Leu Pro Ser His Tyr
290 295 300
Asn Tyr Leu Gln Phe Ala Tyr Tyr Asn Lys Thr Ile Cys Tyr Cys Asn
305 310 315 320
Leu Pro Cys Leu Leu Lys Ile Tyr Arg Lys Lys Lys Ser Ala Lys Glu
325 330 335
Tyr Arg Gln Arg Ser Leu Leu Glu Lys Glu Leu Leu Phe Phe Ala Tyr
340 345 350
Asp Val Phe Gly Ile Pro Phe Val Asp Pro Asp Ser Trp Thr Pro Glu
355 360 365
Glu Val Ile Pro Lys Arg Leu Gln Glu Lys Gln Lys Ser Glu Arg Glu
370 375 380
Thr Ala Val Arg Ile Ser Gln Glu Ile Gly Asn Leu Met Lys Glu Ile
385 390 395 400
Glu Thr Leu Val Glu Glu Lys Thr Lys Glu Ser Leu Asp Val Ser Arg
405 410 415
Leu Thr Arg Glu Gly Gly Pro Leu Leu Tyr Glu Gly Ile Ser Leu Thr
420 425 430
Met Asn Ser Lys Leu Leu Asn Gly Ser Gln Arg Val Val Met Asp Gly
435 440 445
Val Ile Ser Asp His Glu Cys Gln Glu Leu Gln Arg Leu Thr Asn Val
450 455 460
Ala Ala Thr Ser Gly Asp Gly Tyr Arg Gly Gln Thr Ser Pro His Thr
465 470 475 480
Pro Asn Glu Lys Phe Tyr Gly Val Thr Val Phe Lys Ala Leu Lys Leu
485 490 495
Gly Gln Glu Gly Lys Val Pro Leu Gln Ser Ala His Leu Tyr Tyr Asn

```
            500                     505                     510
Val Thr Glu Lys Val Arg Arg Ile Met Glu Ser Tyr Phe Arg Leu Asp
        515                     520                     525
Thr Pro Leu Tyr Phe Ser Tyr Ser His Leu Val Cys Arg Thr Ala Ile
    530                     535                     540
Glu Glu Val Gln Ala Glu Arg Lys Asp Asp Ser His Pro Val His Val
545                     550                     555                 560
Asp Asn Cys Ile Leu Asn Ala Glu Thr Leu Val Cys Val Lys Glu Pro
                565                     570                     575
Pro Ala Tyr Thr Phe Arg Asp Tyr Ser Ala Ile Leu Tyr Leu Asn Gly
            580                     585                     590
Asp Phe Asp Gly Gly Asn Phe Tyr Phe Thr Glu Leu Asp Ala Lys Thr
        595                     600                     605
Val Thr Ala Glu Val Gln Pro Gln Cys Gly Arg Ala Val Gly Phe Ser
    610                     615                     620
Ser Gly Thr Glu Asn Pro His Gly Val Lys Ala Val Thr Arg Gly Gln
625                     630                     635                 640
Arg Cys Ala Ile Ala Leu Trp Phe Thr Leu Asp Pro Arg His Ser Glu
                645                     650                     655
Arg Asp Arg Val Gln Ala Asp Asp Leu Val Lys Met Leu Phe Ser Pro
            660                     665                     670
Glu Glu Met Asp Leu Ser Gln Glu Gln Pro Leu Asp Ala Gln Gln Gly
        675                     680                     685
Pro Pro Glu Pro Ala Gln Glu Ser Leu Ser Gly Ser Glu Ser Lys Pro
    690                     695                     700
Lys Asp Glu Leu
705
```

<210> 26
<211> 736
<212> PRT
<213> Homo sapiens

<400> 26

```
Met Ala Val Arg Ala Leu Lys Leu Leu Thr Thr Leu Leu Ala Val Val
1               5                   10                  15
Ala Ala Ala Ser Gln Ala Glu Val Glu Ser Glu Ala Gly Trp Gly Met
            20                  25                  30
Val Thr Pro Asp Leu Leu Phe Ala Glu Gly Thr Ala Ala Tyr Ala Arg
        35                  40                  45
Gly Asp Trp Pro Gly Val Val Leu Ser Met Glu Arg Ala Leu Arg Ser
    50                  55                  60
```

Arg Ala Ala Leu Arg Ala Leu Arg Leu Arg Cys Arg Thr Gln Cys Ala
65 70 75 80
Ala Asp Phe Pro Trp Glu Leu Asp Pro Asp Trp Ser Pro Ser Pro Ala
85 90 95
Gln Ala Ser Gly Ala Ala Ala Leu Arg Asp Leu Ser Phe Phe Gly Gly
100 105 110
Leu Leu Arg Arg Ala Ala Cys Leu Arg Arg Cys Leu Gly Pro Pro Ala
115 120 125
Ala His Ser Leu Ser Glu Glu Met Glu Leu Glu Phe Arg Lys Arg Ser
130 135 140
Pro Tyr Asn Tyr Leu Gln Val Ala Tyr Phe Lys Ile Asn Lys Leu Glu
145 150 155 160
Lys Ala Val Ala Ala Ala His Thr Phe Phe Val Gly Asn Pro Glu His
165 170 175
Met Glu Met Gln Gln Asn Leu Asp Tyr Tyr Gln Thr Met Ser Gly Val
180 185 190
Lys Glu Ala Asp Phe Lys Asp Leu Glu Thr Gln Pro His Met Gln Glu
195 200 205
Phe Arg Leu Gly Val Arg Leu Tyr Ser Glu Glu Gln Pro Gln Glu Ala
210 215 220
Val Pro His Leu Glu Ala Ala Leu Gln Glu Tyr Phe Val Ala Tyr Glu
225 230 235 240
Glu Cys Arg Ala Leu Cys Glu Gly Pro Tyr Asp Tyr Asp Gly Tyr Asn
245 250 255
Tyr Leu Glu Tyr Asn Ala Asp Leu Phe Gln Ala Ile Thr Asp His Tyr
260 265 270
Ile Gln Val Leu Asn Cys Lys Gln Asn Cys Val Thr Glu Leu Ala Ser
275 280 285
His Pro Ser Arg Glu Lys Pro Phe Glu Asp Phe Leu Pro Ser His Tyr
290 295 300
Asn Tyr Leu Gln Phe Ala Tyr Tyr Asn Ile Gly Asn Tyr Thr Gln Ala
305 310 315 320
Val Glu Cys Ala Lys Thr Tyr Leu Leu Phe Phe Pro Asn Asp Glu Val
325 330 335
Met Asn Gln Asn Leu Ala Tyr Tyr Ala Ala Met Leu Gly Glu Glu His
340 345 350
Thr Arg Ser Ile Gly Pro Arg Glu Ser Ala Lys Glu Tyr Arg Gln Arg
355 360 365
Ser Leu Leu Glu Lys Glu Leu Leu Phe Phe Ala Tyr Asp Val Phe Gly
370 375 380
Ile Pro Phe Val Asp Pro Asp Ser Trp Thr Pro Glu Glu Val Ile Pro
385 390 395 400

Lys Arg Leu Gln Glu Lys Gln Lys Ser Glu Arg Glu Thr Ala Val Arg
405 410 415

Ile Ser Gln Glu Ile Gly Asn Leu Met Lys Glu Ile Glu Thr Leu Val
420 425 430

Glu Glu Lys Thr Lys Glu Ser Leu Asp Val Ser Arg Leu Thr Arg Glu
435 440 445

Gly Gly Pro Leu Leu Tyr Glu Gly Ile Ser Leu Thr Met Asn Ser Lys
450 455 460

Leu Leu Asn Gly Ser Gln Arg Val Val Met Asp Gly Val Ile Ser Asp
465 470 475 480

His Glu Cys Gln Glu Leu Gln Arg Leu Thr Asn Val Ala Ala Thr Ser
485 490 495

Gly Asp Gly Tyr Arg Gly Gln Thr Ser Pro His Thr Pro Asn Glu Lys
500 505 510

Phe Tyr Gly Val Thr Val Phe Lys Ala Leu Lys Leu Gly Gln Glu Gly
515 520 525

Lys Val Pro Leu Gln Ser Ala His Leu Tyr Tyr Asn Val Thr Glu Lys
530 535 540

Val Arg Arg Ile Met Glu Ser Tyr Phe Arg Leu Asp Thr Pro Leu Tyr
545 550 555 560

Phe Ser Tyr Ser His Leu Val Cys Arg Thr Ala Ile Glu Glu Val Gln
565 570 575

Ala Glu Arg Lys Asp Asp Ser His Pro Val His Val Asp Asn Cys Ile
580 585 590

Leu Asn Ala Glu Thr Leu Val Cys Val Lys Glu Pro Pro Ala Tyr Thr
595 600 605

Phe Arg Asp Tyr Ser Ala Ile Leu Tyr Leu Asn Gly Asp Phe Asp Gly
610 615 620

Gly Asn Phe Tyr Phe Thr Glu Leu Asp Ala Lys Thr Val Thr Ala Glu
625 630 635 640

Val Gln Pro Gln Cys Gly Arg Ala Val Gly Phe Ser Ser Gly Thr Glu
645 650 655

Asn Pro His Gly Val Lys Ala Val Thr Arg Gly Gln Arg Cys Ala Ile
660 665 670

Ala Leu Trp Phe Thr Leu Asp Pro Arg His Ser Glu Arg Asp Arg Val
675 680 685

Gln Ala Asp Asp Leu Val Lys Met Leu Phe Ser Pro Glu Glu Met Asp
690 695 700

Leu Ser Gln Glu Gln Pro Leu Asp Ala Gln Gln Gly Pro Pro Glu Pro
705 710 715 720

Ala Gln Glu Ser Leu Ser Gly Ser Glu Ser Lys Pro Lys Asp Glu Leu
725 730 735

<210> 27
<211> 478
<212> PRT
<213> Homo sapiens

<400> 27
Met Ser Pro Pro Leu Leu Lys Leu Gly Ala Val Leu Ser Thr Met Ala
1 5 10 15
Met Ile Ser Asn Trp Met Ser Gln Thr Leu Pro Ser Leu Val Gly Leu
20 25 30
Asn Thr Thr Arg Leu Ser Thr Pro Asp Thr Leu Thr Gln Ile Ser Pro
35 40 45
Lys Glu Gly Trp Gln Val Tyr Ser Ser Ala Gln Asp Pro Asp Gly Arg
50 55 60
Cys Ile Cys Thr Val Val Ala Pro Glu Gln Asn Leu Cys Ser Arg Asp
65 70 75 80
Ala Lys Ser Arg Gln Leu Arg Gln Leu Leu Glu Lys Val Gln Asn Met
85 90 95
Ser Gln Ser Ile Glu Val Leu Asn Leu Arg Thr Gln Arg Asp Phe Gln
100 105 110
Tyr Val Leu Lys Met Glu Thr Gln Met Lys Gly Leu Lys Ala Lys Phe
115 120 125
Arg Gln Ile Glu Asp Asp Arg Lys Thr Leu Met Thr Lys His Phe Gln
130 135 140
Glu Leu Lys Glu Lys Met Asp Glu Leu Leu Pro Leu Ile Pro Val Leu
145 150 155 160
Glu Gln Tyr Lys Thr Asp Ala Lys Leu Ile Thr Gln Phe Lys Glu Glu
165 170 175
Ile Arg Asn Leu Ser Ala Val Leu Thr Gly Ile Gln Glu Glu Ile Gly
180 185 190
Ala Tyr Asp Tyr Glu Glu Leu His Gln Arg Val Leu Ser Leu Glu Thr
195 200 205
Arg Leu Arg Asp Cys Met Lys Lys Leu Thr Cys Gly Lys Leu Met Lys
210 215 220
Ile Thr Gly Pro Val Thr Val Lys Thr Ser Gly Thr Arg Phe Gly Ala
225 230 235 240
Trp Met Thr Asp Pro Leu Ala Ser Glu Lys Asn Asn Arg Val Trp Tyr
245 250 255
Met Asp Ser Tyr Thr Asn Asn Lys Ile Val Arg Glu Tyr Lys Ser Ile
260 265 270
Ala Asp Phe Val Ser Gly Ala Glu Ser Arg Thr Tyr Asn Leu Pro Phe

275 280 285
Lys Trp Ala Gly Thr Asn His Val Val Tyr Asn Gly Ser Leu Tyr Phe
290 295 300
Asn Lys Tyr Gln Ser Asn Ile Ile Ile Lys Tyr Ser Phe Asp Met Gly
305 310 315 320
Arg Val Leu Ala Gln Arg Ser Leu Glu Tyr Ala Gly Phe His Asn Val
325 330 335
Tyr Pro Tyr Thr Trp Gly Gly Phe Ser Asp Ile Asp Leu Met Ala Asp
340 345 350
Glu Ile Gly Leu Trp Ala Val Tyr Ala Thr Asn Gln Asn Ala Gly Asn
355 360 365
Ile Val Ile Ser Gln Leu Asn Gln Asp Thr Leu Glu Val Met Lys Ser
370 375 380
Trp Ser Thr Gly Tyr Pro Lys Arg Ser Ala Gly Glu Ser Phe Met Ile
385 390 395 400
Cys Gly Thr Leu Tyr Val Thr Asn Ser His Leu Thr Gly Ala Lys Val
405 410 415
Tyr Tyr Ser Tyr Ser Thr Lys Thr Ser Thr Tyr Glu Tyr Thr Asp Ile
420 425 430
Pro Phe His Asn Gln Tyr Phe His Ile Ser Met Leu Asp Tyr Asn Ala
435 440 445
Arg Asp Arg Ala Leu Tyr Ala Trp Asn Asn Gly His Gln Val Leu Phe
450 455 460
Asn Val Thr Leu Phe His Ile Ile Lys Thr Glu Asp Asp Thr
465 470 475

<210> 28
<211> 589
<212> PRT
<213> Homo sapiens

<400> 28
Met Trp Thr Ser Gly Arg Met Ser Asn Ala Lys Asn Trp Leu Gly Leu
1 5 10 15
Gly Met Ser Leu Tyr Phe Trp Gly Leu Met Asp Leu Thr Thr Thr Val
20 25 30
Leu Ser Asp Thr Pro Thr Pro Gln Gly Glu Leu Glu Ala Leu Leu Ser
35 40 45
Asp Lys Pro Gln Ser His Gln Arg Thr Lys Arg Ser Trp Val Trp Asn
50 55 60
Gln Phe Phe Val Leu Glu Glu Tyr Thr Gly Thr Asp Pro Leu Tyr Val
65 70 75 80

Gly Lys Leu His Ser Asp Met Asp Arg Gly Asp Gly Ser Ile Lys Tyr
85 90 95
Ile Leu Ser Gly Glu Gly Ala Gly Ile Val Phe Thr Ile Asp Asp Thr
100 105 110
Thr Gly Asp Ile His Ala Ile Gln Arg Leu Asp Arg Glu Glu Arg Ala
115 120 125
Gln Tyr Thr Leu Arg Ala Gln Ala Leu Asp Arg Arg Thr Gly Arg Pro
130 135 140
Met Glu Pro Glu Ser Glu Phe Ile Ile Lys Ile Gln Asp Ile Asn Asp
145 150 155 160
Asn Glu Pro Lys Phe Leu Asp Gly Pro Tyr Val Ala Thr Val Pro Glu
165 170 175
Met Ser Pro Val Gly Thr Ser Val Ile Gln Val Thr Ala Thr Asp Ala
180 185 190
Asp Asp Pro Thr Tyr Gly Asn Ser Ala Arg Val Val Tyr Ser Ile Leu
195 200 205
Gln Gly Gln Pro Tyr Phe Ser Val Asp Ser Lys Thr Gly Val Ile Arg
210 215 220
Thr Ala Leu Met Asn Met Asp Arg Glu Ala Lys Glu Tyr Tyr Glu Val
225 230 235 240
Ile Ile Gln Ala Lys Asp Met Gly Gly Gln Leu Gly Gly Leu Ala Gly
245 250 255
Thr Thr Thr Val Asn Ile Thr Leu Ser Asp Val Asn Asp Asn Pro Pro
260 265 270
Arg Phe Pro Gln Lys His Tyr Gln Met Ser Val Leu Glu Ser Ala Pro
275 280 285
Ile Ser Ser Thr Val Gly Arg Val Phe Ala Lys Asp Leu Asp Glu Gly
290 295 300
Ile Asn Ala Glu Met Lys Tyr Thr Ile Val Asp Gly Asp Gly Ala Asp
305 310 315 320
Ala Phe Asp Ile Ser Thr Asp Pro Asn Phe Gln Val Gly Ile Ile Thr
325 330 335
Val Lys Lys Pro Leu Ser Phe Glu Ser Lys Lys Ser Tyr Thr Leu Lys
340 345 350
Val Glu Gly Ala Asn Pro His Leu Glu Met Arg Phe Leu Asn Leu Gly
355 360 365
Pro Phe Gln Asp Thr Thr Thr Val His Ile Ser Val Glu Asp Val Asp
370 375 380
Glu Pro Pro Val Phe Glu Pro Gly Phe Tyr Phe Val Glu Val Pro Glu
385 390 395 400
Asp Val Ala Ile Gly Thr Thr Ile Gln Ile Ile Ser Ala Lys Asp Pro
405 410 415

Asp Val Thr Asn Asn Ser Ile Arg Tyr Ser Ile Asp Arg Ser Ser Asp
420 425 430
Pro Gly Arg Phe Phe Tyr Val Asp Ile Thr Thr Gly Ala Leu Met Thr
435 440 445
Ala Arg Pro Leu Asp Arg Glu Glu Phe Ser Trp His Asn Ile Thr Val
450 455 460
Leu Ala Met Glu Met Asn Asn Pro Ser Gln Val Gly Ser Val Pro Val
465 470 475 480
Thr Ile Lys Val Leu Asp Val Asn Asp Asn Ala Pro Glu Phe Pro Arg
485 490 495
Phe Tyr Glu Ala Phe Val Cys Glu Asn Ala Lys Ala Gly Gln Leu Ile
500 505 510
Gln Thr Val Ser Ala Val Asp Gln Asp Asp Pro Arg Asn Gly Gln His
515 520 525
Phe Tyr Tyr Ser Leu Ala Pro Glu Ala Ala Asn Asn Pro Asn Phe Thr
530 535 540
Ile Arg Asp Asn Gln Gly Asn Gln Val Asp Gly Trp Leu Ser Val Leu
545 550 555 560
Phe Tyr Ser Ile Gly Gln Leu Leu Trp Val Thr Val Leu Cys Lys Gln
565 570 575
Cys Gln Arg Leu Pro Val Pro Tyr Gln Gln Gly Gly Cys
580 585

<210> 29
<211> 801
<212> PRT
<213> Homo sapiens

<400> 29
Met Trp Thr Ser Gly Arg Met Ser Asn Ala Lys Asn Trp Leu Gly Leu
1 5 10 15
Gly Met Ser Leu Tyr Phe Trp Gly Leu Met Asp Leu Thr Thr Thr Val
20 25 30
Leu Ser Asp Thr Pro Thr Pro Gln Gly Glu Leu Glu Ala Leu Leu Ser
35 40 45
Asp Lys Pro Gln Ser His Gln Arg Thr Lys Arg Ser Trp Val Trp Asn
50 55 60
Gln Phe Phe Val Leu Glu Glu Tyr Thr Gly Thr Asp Pro Leu Tyr Val
65 70 75 80
Gly Lys Leu His Ser Asp Met Asp Arg Gly Asp Gly Ser Ile Lys Tyr
85 90 95
Ile Leu Ser Gly Glu Gly Ala Gly Ile Val Phe Thr Ile Asp Asp Thr

100 105 110

Thr Gly Asp Ile His Ala Ile Gln Arg Leu Asp Arg Glu Glu Arg Ala

115 120 125

Gln Tyr Thr Leu Arg Ala Gln Ala Leu Asp Arg Arg Thr Gly Arg Pro

130 135 140

Met Glu Pro Glu Ser Glu Phe Ile Ile Lys Ile Gln Asp Ile Asn Asp

145 150 155 160

Asn Glu Pro Lys Phe Leu Asp Gly Pro Tyr Val Ala Thr Val Pro Glu

165 170 175

Met Ser Pro Val Gly Thr Ser Val Ile Gln Val Thr Ala Thr Asp Ala

180 185 190

Asp Asp Pro Thr Tyr Gly Asn Ser Ala Arg Val Val Tyr Ser Ile Leu

195 200 205

Gln Gly Gln Pro Tyr Phe Ser Val Asp Ser Lys Thr Gly Val Ile Arg

210 215 220

Thr Ala Leu Met Asn Met Asp Arg Glu Ala Lys Glu Tyr Tyr Glu Val

225 230 235 240

Ile Ile Gln Ala Lys Asp Met Gly Gly Gln Leu Gly Gly Leu Ala Gly

245 250 255

Thr Thr Thr Val Asn Ile Thr Leu Ser Asp Val Asn Asp Asn Pro Pro

260 265 270

Arg Phe Pro Gln Lys His Tyr Gln Met Ser Val Leu Glu Ser Ala Pro

275 280 285

Ile Ser Ser Thr Val Gly Arg Val Phe Ala Lys Asp Leu Asp Glu Gly

290 295 300

Ile Asn Ala Glu Met Lys Tyr Thr Ile Val Asp Gly Asp Gly Ala Asp

305 310 315 320

Ala Phe Asp Ile Ser Thr Asp Pro Asn Phe Gln Val Gly Ile Ile Thr

325 330 335

Val Lys Lys Pro Leu Ser Phe Glu Ser Lys Lys Ser Tyr Thr Leu Lys

340 345 350

Val Glu Gly Ala Asn Pro His Leu Glu Met Arg Phe Leu Asn Leu Gly

355 360 365

Pro Phe Gln Asp Thr Thr Thr Val His Ile Ser Val Glu Asp Val Asp

370 375 380

Glu Pro Pro Val Phe Glu Pro Gly Phe Tyr Phe Val Glu Val Pro Glu

385 390 395 400

Asp Val Ala Ile Gly Thr Thr Ile Gln Ile Ile Ser Ala Lys Asp Pro

405 410 415

Asp Val Thr Asn Asn Ser Ile Arg Tyr Ser Ile Asp Arg Ser Ser Asp

420 425 430

Pro Gly Arg Phe Phe Tyr Val Asp Ile Thr Thr Gly Ala Leu Met Thr

435 440 445

Ala Arg Pro Leu Asp Arg Glu Glu Phe Ser Trp His Asn Ile Thr Val

450 455 460

Leu Ala Met Glu Met Asn Asn Pro Ser Gln Val Gly Ser Val Pro Val

465 470 475 480

Thr Ile Lys Val Leu Asp Val Asn Asp Asn Ala Pro Glu Phe Pro Arg

485 490 495

Phe Tyr Glu Ala Phe Val Cys Glu Asn Ala Lys Ala Gly Gln Leu Ile

500 505 510

Gln Thr Val Ser Ala Val Asp Gln Asp Asp Pro Arg Asn Gly Gln His

515 520 525

Phe Tyr Tyr Ser Leu Ala Pro Glu Ala Ala Asn Asn Pro Asn Phe Thr

530 535 540

Ile Arg Asp Asn Gln Asp Asn Thr Ala Arg Ile Leu Thr Arg Arg Ser

545 550 555 560

Gly Phe Arg Gln Gln Glu Gln Ser Val Phe His Leu Pro Ile Leu Ile

565 570 575

Ala Asp Ser Gly Gln Pro Val Leu Ser Ser Thr Gly Thr Leu Thr Ile

580 585 590

Gln Val Cys Ser Cys Asp Asp Asp Gly His Val Met Ser Cys Ser Pro

595 600 605

Glu Ala Tyr Met Leu Pro Val Ser Leu Ser Arg Gly Ala Leu Ile Ala

610 615 620

Ile Leu Ala Cys Ile Phe Val Leu Leu Val Leu Val Leu Leu Ile Leu

625 630 635 640

Ser Met Arg Arg His Arg Lys Gln Pro Tyr Ile Ile Asp Asp Glu Glu

645 650 655

Asn Ile His Glu Asn Ile Val Arg Tyr Asp Asp Glu Gly Gly Gly Glu

660 665 670

Glu Asp Thr Glu Ala Phe Asp Ile Ala Ala Met Trp Asn Pro Arg Glu

675 680 685

Ala Gln Ala Gly Ala Ala Pro Lys Thr Arg Gln Asp Met Leu Pro Glu

690 695 700

Ile Glu Ser Leu Ser Arg Tyr Val Pro Gln Thr Cys Ala Val Asn Ser

705 710 715 720

Thr Val His Ser Tyr Val Leu Ala Lys Leu Tyr Glu Ala Asp Met Asp

725 730 735

Leu Trp Ala Pro Pro Phe Asp Ser Leu Gln Thr Tyr Met Phe Glu Gly

740 745 750

Asp Gly Ser Val Ala Gly Ser Leu Ser Ser Leu Gln Ser Ala Thr Ser

755 760 765

Asp Ser Glu Gln Ser Phe Asp Phe Leu Thr Asp Trp Gly Pro Arg Phe

```
    770                 775                 780
Arg Lys Leu Ala Glu Leu Tyr Gly Ala Ser Glu Gly Pro Ala Pro Leu
785                 790                 795                 800
Trp
```

```
      <210> 30
      <211> 287
      <212> PRT
      <213> Homo sapiens

      <400> 30
Met Ala Leu Gly Leu Leu Ile Ala Val Pro Leu Leu Leu Gln Ala Ala
 1               5                  10                  15
Pro Pro Gly Ala Ala His Tyr Glu Met Leu Gly Thr Cys Arg Met Ile
            20                  25                  30
Cys Asp Pro Tyr Ser Val Ala Pro Ala Gly Gly Pro Ala Gly Ala Lys
        35                  40                  45
Ala Pro Pro Pro Gly Pro Ser Thr Ala Ala Leu Glu Val Met Gln Asp
    50                  55                  60
Leu Ser Ala Asn Pro Pro Pro Pro Phe Ile Gln Gly Pro Lys Gly Asp
65                  70                  75                  80
Pro Gly Arg Pro Gly Lys Pro Gly Pro Arg Gly Pro Pro Gly Glu Pro
                85                  90                  95
Gly Pro Pro Gly Pro Arg Gly Pro Pro Gly Glu Lys Gly Asp Ser Gly
            100                 105                 110
Arg Pro Gly Leu Pro Gly Leu Gln Leu Thr Thr Ser Ala Ala Gly Gly
        115                 120                 125
Val Gly Val Val Ser Gly Gly Thr Gly Gly Gly Gly Asp Thr Glu Gly
    130                 135                 140
Glu Val Thr Ser Ala Leu Ser Ala Ala Phe Ser Gly Pro Lys Ile Ala
145                 150                 155                 160
Phe Tyr Val Gly Leu Lys Ser Pro His Glu Gly Tyr Glu Val Leu Lys
                165                 170                 175
Phe Asp Asp Val Val Thr Asn Leu Gly Asn His Tyr Asp Pro Thr Thr
            180                 185                 190
Gly Lys Phe Ser Cys Gln Val Arg Gly Ile Tyr Phe Phe Thr Tyr His
        195                 200                 205
Ile Leu Met Arg Gly Gly Asp Gly Thr Ser Met Trp Ala Asp Leu Cys
    210                 215                 220
Lys Asn Gly Gln Val Arg Ala Ser Ala Ile Ala Gln Asp Ala Asp Gln
225                 230                 235                 240
```

Asn Tyr Asp Tyr Ala Ser Asn Ser Val Val Leu His Leu Asp Ser Gly
245 250 255
Asp Glu Val Tyr Val Lys Leu Asp Gly Gly Lys Ala His Gly Gly Asn
260 265 270
Asn Asn Lys Tyr Ser Thr Phe Ser Gly Phe Leu Leu Tyr Pro Asp
275 280 285

<210> 31
<211> 159
<212> PRT
<213> Homo sapiens

<400> 31
Met Lys Ala Trp Gly Thr Val Val Val Thr Leu Ala Thr Leu Met Val
1 5 10 15
Val Thr Val Asp Ala Lys Ile Tyr Glu Arg Cys Glu Leu Ala Ala Arg
20 25 30
Leu Glu Arg Ala Gly Leu Asn Gly Tyr Lys Gly Tyr Gly Val Gly Asp
35 40 45
Trp Leu Cys Met Ala His Tyr Glu Ser Gly Phe Asp Thr Ala Phe Val
50 55 60
Asp His Asn Pro Asp Gly Ser Ser Glu Tyr Gly Ile Phe Gln Leu Asn
65 70 75 80
Ser Ala Trp Trp Cys Asp Asn Gly Ile Thr Pro Thr Lys Asn Leu Cys
85 90 95
His Met Asp Cys His Asp Leu Leu Asn Arg His Ile Leu Asp Asp Ile
100 105 110
Arg Cys Ala Lys Gln Ile Val Ser Ser Gln Asn Gly Leu Ser Ala Trp
115 120 125
Thr Ser Trp Arg Leu His Cys Ser Gly His Asp Leu Ser Glu Trp Leu
130 135 140
Lys Gly Cys Asp Met His Val Lys Ile Asp Pro Lys Ile His Pro
145 150 155

<210> 32
<211> 220
<212> PRT
<213> Homo sapiens

<400> 32
Met Val Arg Asn Ile Phe Lys Thr Phe Pro Ser Val Phe Thr Gly Asn
1 5 10 15

Val Val Ser Gln Ser Ser Leu Thr Pro Leu Met Val Asn Gly Ile Leu
20 25 30
Gly Glu Ser Val Thr Leu Pro Leu Glu Phe Pro Ala Gly Glu Lys Val
35 40 45
Asn Phe Ile Thr Trp Leu Phe Asn Glu Thr Ser Leu Ala Phe Ile Val
50 55 60
Pro His Glu Thr Lys Ser Pro Glu Ile His Val Thr Asn Pro Lys Gln
65 70 75 80
Gly Lys Arg Leu Asn Phe Thr Gln Ser Tyr Ser Leu Gln Leu Ser Asn
85 90 95
Leu Lys Met Glu Asp Thr Gly Ser Tyr Arg Ala Gln Ile Ser Thr Lys
100 105 110
Thr Ser Ala Lys Leu Ser Ser Tyr Thr Leu Arg Ile Leu Arg Gln Leu
115 120 125
Arg Asn Ile Gln Val Thr Asn His Ser Gln Leu Phe Gln Asn Met Thr
130 135 140
Cys Glu Leu His Leu Thr Cys Ser Val Glu Asp Ala Asp Asp Asn Val
145 150 155 160
Ser Phe Arg Trp Glu Ala Leu Gly Asn Thr Leu Ser Ser Gln Pro Asn
165 170 175
Leu Thr Val Ser Trp Asp Pro Arg Ile Ser Ser Glu Gln Asp Tyr Thr
180 185 190
Cys Ile Ala Glu Asn Ala Val Ser Asn Leu Ser Phe Ser Val Ser Ala
195 200 205
Gln Lys Leu Cys Glu Gly Asn Ser Leu Pro Gln Val
210 215 220

<210> 33
<211> 346
<212> PRT
<213> Homo sapiens

<400> 33
Met Thr Ala Ser Arg Ser Gln Ala Pro Val Phe Thr Ala Glu Ser Met
1 5 10 15
Leu Trp Leu Phe Gln Ser Leu Leu Phe Val Phe Cys Phe Gly Pro Gly
20 25 30
Asn Val Val Ser Gln Ser Ser Leu Thr Pro Leu Met Val Asn Gly Ile
35 40 45
Leu Gly Glu Ser Val Thr Leu Pro Leu Glu Phe Pro Ala Gly Glu Lys
50 55 60
Val Asn Phe Ile Thr Trp Leu Phe Asn Glu Thr Ser Leu Ala Phe Ile

65 70 75 80
Val Pro His Glu Thr Lys Ser Pro Glu Ile His Val Thr Asn Pro Lys
85 90 95
Gln Gly Lys Arg Leu Asn Phe Thr Gln Ser Tyr Ser Leu Gln Leu Ser
100 105 110
Asn Leu Lys Met Glu Asp Thr Gly Ser Tyr Arg Ala Gln Ile Ser Thr
115 120 125
Lys Thr Ser Ala Lys Leu Ser Ser Tyr Thr Leu Arg Ile Leu Arg Gln
130 135 140
Leu Arg Asn Ile Gln Val Thr Asn His Ser Gln Leu Phe Gln Asn Met
145 150 155 160
Thr Cys Glu Leu His Leu Thr Cys Ser Val Glu Asp Ala Asp Asp Asn
165 170 175
Val Ser Phe Arg Trp Glu Ala Leu Gly Asn Thr Leu Ser Ser Gln Pro
180 185 190
Asn Leu Thr Val Ser Trp Asp Pro Arg Ile Ser Ser Glu Gln Asp Tyr
195 200 205
Thr Cys Ile Ala Glu Asn Ala Val Ser Asn Leu Ser Phe Ser Val Ser
210 215 220
Ala Gln Lys Leu Cys Glu Asp Val Lys Ile Gln Tyr Thr Asp Thr Lys
225 230 235 240
Met Ile Leu Phe Met Val Ser Gly Ile Cys Ile Val Phe Gly Phe Ile
245 250 255
Ile Leu Leu Leu Leu Val Leu Arg Lys Arg Arg Asp Ser Leu Ser Leu
260 265 270
Ser Thr Gln Arg Thr Gln Gly Pro Glu Ser Ala Arg Asn Leu Glu Tyr
275 280 285
Val Ser Val Ser Pro Thr Asn Asn Thr Val Tyr Ala Ser Val Thr His
290 295 300
Ser Asn Arg Glu Thr Glu Ile Trp Thr Pro Arg Glu Asn Asp Thr Ile
305 310 315 320
Thr Ile Tyr Ser Thr Ile Asn His Ser Lys Glu Ser Lys Pro Thr Phe
325 330 335
Ser Arg Ala Thr Ala Leu Asp Asn Val Val
340 345

<210> 34
<211> 1075
<212> PRT
<213> Homo sapiens

<400> 34

```
Met Gly Thr Ala Tyr Leu Cys Cys Pro Gln Val Leu Leu Leu Leu Cys
 1               5                  10                  15
Leu Pro Arg Arg Val Lys Leu Trp Ala Asp Thr Phe Gly Gly Asp Leu
            20                  25                  30
Tyr Asn Thr Val Thr Lys Tyr Ser Gly Ser Leu Leu Leu Gln Lys Lys
        35                  40                  45
Tyr Lys Asp Val Glu Ser Ser Leu Lys Ile Glu Glu Val Asp Gly Leu
    50                  55                  60
Glu Leu Val Arg Lys Phe Ser Glu Asp Met Glu Asn Met Leu Arg Arg
65                  70                  75                  80
Lys Val Glu Ala Val Gln Asn Leu Val Glu Ala Ala Glu Glu Ala Asp
                85                  90                  95
Leu Asn His Glu Phe Asn Glu Ser Leu Val Phe Asp Tyr Tyr Asn Ser
            100                 105                 110
Val Leu Ile Asn Glu Arg Asp Glu Lys Gly Asn Phe Val Glu Leu Gly
        115                 120                 125
Ala Glu Phe Leu Leu Glu Ser Asn Ala His Phe Ser Asn Leu Pro Val
    130                 135                 140
Asn Thr Ser Ile Ser Ser Val Gln Leu Pro Thr Asn Val Tyr Asn Lys
145                 150                 155                 160
Asp Pro Asp Ile Leu Asn Gly Val Tyr Met Ser Glu Ala Leu Asn Ala
                165                 170                 175
Val Phe Val Glu Asn Phe Gln Arg Asp Pro Thr Leu Thr Trp Gln Tyr
            180                 185                 190
Phe Gly Ser Ala Thr Gly Phe Phe Arg Ile Tyr Pro Gly Ile Lys Trp
        195                 200                 205
Thr Pro Asp Glu Asn Gly Val Ile Thr Phe Asp Cys Arg Asn Arg Gly
    210                 215                 220
Trp Tyr Ile Gln Ala Ala Thr Ser Pro Lys Asp Ile Val Ile Leu Val
225                 230                 235                 240
Asp Val Ser Gly Ser Met Lys Gly Leu Arg Met Thr Ile Ala Lys His
                245                 250                 255
Thr Ile Thr Thr Ile Leu Asp Thr Leu Gly Glu Asn Asp Phe Ile Asn
            260                 265                 270
Ile Ile Ala Tyr Asn Asp Tyr Val His Tyr Ile Glu Pro Cys Phe Lys
        275                 280                 285
Gly Ile Leu Val Gln Ala Asp Arg Asp Asn Arg Glu His Phe Lys Leu
    290                 295                 300
Leu Val Glu Glu Leu Met Val Lys Gly Val Gly Val Val Asp Gln Ala
305                 310                 315                 320
Leu Arg Glu Ala Phe Gln Ile Leu Lys Gln Phe Gln Glu Ala Lys Gln
                325                 330                 335
```

Gly Ser Leu Cys Asn Gln Ala Ile Met Leu Ile Ser Asp Gly Ala Val
340 345 350

Glu Asp Tyr Glu Pro Val Phe Glu Lys Tyr Asn Trp Pro Asp Cys Lys
355 360 365

Val Arg Val Phe Thr Tyr Leu Ile Gly Arg Glu Val Ser Phe Ala Asp
370 375 380

Arg Met Lys Trp Ile Ala Cys Asn Asn Lys Gly Tyr Tyr Thr Gln Ile
385 390 395 400

Ser Thr Leu Ala Asp Thr Gln Glu Asn Val Met Glu Tyr Leu His Val
405 410 415

Leu Ser Arg Pro Met Val Ile Asn His Asp His Asp Ile Ile Trp Thr
420 425 430

Glu Ala Tyr Met Asp Ser Lys Leu Leu Ser Ser Gln Ala Gln Ser Leu
435 440 445

Thr Leu Leu Thr Thr Val Ala Met Pro Val Phe Ser Lys Lys Asn Glu
450 455 460

Thr Arg Ser His Gly Ile Leu Leu Gly Val Val Gly Ser Asp Val Ala
465 470 475 480

Leu Arg Glu Leu Met Lys Leu Ala Pro Arg Tyr Lys Leu Gly Val His
485 490 495

Gly Tyr Ala Phe Leu Asn Thr Asn Asn Gly Tyr Ile Leu Ser His Pro
500 505 510

Asp Leu Arg Pro Leu Tyr Arg Glu Gly Lys Lys Leu Lys Pro Lys Pro
515 520 525

Asn Tyr Asn Ser Val Asp Leu Ser Glu Val Glu Trp Glu Asp Gln Ala
530 535 540

Glu Ser Leu Arg Thr Ala Met Ile Asn Arg Glu Thr Gly Thr Leu Ser
545 550 555 560

Met Asp Val Lys Val Pro Met Asp Lys Gly Lys Arg Val Leu Phe Leu
565 570 575

Thr Asn Asp Tyr Phe Phe Thr Asp Ile Ser Asp Thr Pro Phe Ser Leu
580 585 590

Gly Val Val Leu Ser Arg Gly His Gly Glu Tyr Ile Leu Leu Gly Asn
595 600 605

Thr Ser Val Glu Glu Gly Leu His Asp Leu Leu His Pro Asp Leu Ala
610 615 620

Leu Ala Gly Asp Trp Ile Tyr Cys Ile Thr Asp Ile Asp Pro Asp His
625 630 635 640

Arg Lys Leu Ser Gln Leu Glu Ala Met Ile Arg Phe Leu Thr Arg Lys
645 650 655

Asp Pro Asp Leu Glu Cys Asp Glu Glu Leu Val Arg Glu Val Leu Phe
660 665 670

```
Asp Ala Val Val Thr Ala Pro Met Glu Ala Tyr Trp Thr Ala Leu Ala
        675                 680                 685
Leu Asn Met Ser Glu Glu Ser Glu His Val Val Asp Met Ala Phe Leu
    690                 695                 700
Gly Thr Arg Ala Gly Leu Leu Arg Ser Ser Leu Phe Val Gly Ser Glu
705                 710                 715                 720
Lys Val Ser Asp Arg Lys Phe Leu Thr Pro Glu Asp Glu Ala Ser Val
                725                 730                 735
Phe Thr Leu Asp Arg Phe Pro Leu Trp Tyr Arg Gln Ala Ser Glu His
            740                 745                 750
Pro Ala Gly Ser Phe Val Phe Asn Leu Arg Trp Ala Glu Gly Pro Glu
        755                 760                 765
Ser Ala Gly Glu Pro Met Val Val Thr Ala Ser Thr Ala Val Ala Val
    770                 775                 780
Thr Val Asp Lys Arg Thr Ala Ile Ala Ala Ala Ala Gly Val Gln Met
785                 790                 795                 800
Lys Leu Glu Phe Leu Gln Arg Lys Phe Trp Ala Ala Thr Arg Gln Cys
                805                 810                 815
Ser Thr Val Asp Gly Pro Cys Thr Gln Ser Cys Glu Asp Ser Asp Leu
            820                 825                 830
Asp Cys Phe Val Ile Asp Asn Asn Gly Phe Ile Leu Ile Ser Lys Arg
        835                 840                 845
Ser Arg Glu Thr Gly Arg Phe Leu Gly Glu Val Asp Gly Ala Val Leu
    850                 855                 860
Thr Gln Leu Leu Ser Met Gly Val Phe Ser Gln Val Thr Met Tyr Asp
865                 870                 875                 880
Tyr Gln Ala Met Cys Lys Pro Ser Ser His His His Ser Ala Ala Gln
                885                 890                 895
Pro Leu Val Ser Pro Ile Ser Ala Phe Leu Thr Ala Thr Arg Trp Leu
            900                 905                 910
Leu Gln Glu Leu Val Leu Phe Leu Leu Glu Trp Ser Val Trp Gly Ser
        915                 920                 925
Trp Tyr Asp Arg Gly Ala Glu Ala His Lys His Lys Lys Gln Asp Pro
    930                 935                 940
Leu Gln Pro Cys Asp Thr Glu Tyr Pro Val Phe Val Tyr Gln Pro Ala
945                 950                 955                 960
Ile Arg Glu Ala Asn Gly Ile Val Glu Cys Gly Pro Cys Gln Lys Val
                965                 970                 975
Phe Val Val Gln Gln Ile Pro Asn Ser Asn Leu Leu Leu Leu Val Thr
            980                 985                 990
Asp Pro Thr Phe Cys Arg Met Gly Ser Gly Pro Glu Ile Leu Thr Leu
        995                 1000                1005
```

Thr Val Ala Ser Ala His Asn Ala Ser Val Lys Cys Asp Arg Met Arg
1010 1015 1020

Ser Gln Lys Leu Arg Arg Arg Pro Asp Ser Cys His Ala Phe His Pro
1025 1030 1035 1040

Glu Glu Asn Ala Gln Asp Cys Gly Gly Ala Ser Asp Thr Ser Ala Ser
1045 1050 1055

Pro Pro Leu Leu Leu Leu Pro Val Cys Ala Trp Gly Leu Leu Pro Gln
1060 1065 1070

Leu Leu Arg
1075

<210> 35

<211> 1114

<212> PRT

<213> Homo sapiens

<400> 35

Met Pro Ala Thr Pro Asn Phe Leu Ala Asn Pro Ser Ser Ser Ser Arg
1 5 10 15

Trp Ile Pro Leu Gln Pro Met Pro Val Ala Trp Ala Phe Val Gln Lys
20 25 30

Thr Ser Ala Leu Leu Trp Leu Leu Leu Leu Gly Thr Ser Leu Ser Pro
35 40 45

Ala Trp Gly Gln Ala Lys Ile Pro Leu Glu Thr Val Lys Leu Trp Ala
50 55 60

Asp Thr Phe Gly Gly Asp Leu Tyr Asn Thr Val Thr Lys Tyr Ser Gly
65 70 75 80

Ser Leu Leu Leu Gln Lys Lys Tyr Lys Asp Val Glu Ser Ser Leu Lys
85 90 95

Ile Glu Glu Val Asp Gly Leu Glu Leu Val Arg Lys Phe Ser Glu Asp
100 105 110

Met Glu Asn Met Leu Arg Arg Lys Val Glu Ala Val Gln Asn Leu Val
115 120 125

Glu Ala Ala Glu Glu Ala Asp Leu Asn His Glu Phe Asn Glu Ser Leu
130 135 140

Val Phe Asp Tyr Tyr Asn Ser Val Leu Ile Asn Glu Arg Asp Glu Lys
145 150 155 160

Gly Asn Phe Val Glu Leu Gly Ala Glu Phe Leu Leu Glu Ser Asn Ala
165 170 175

His Phe Ser Asn Leu Pro Val Asn Thr Ser Ile Ser Ser Val Gln Leu
180 185 190

Pro Thr Asn Val Tyr Asn Lys Asp Pro Asp Ile Leu Asn Gly Val Tyr

195 200 205
Met Ser Glu Ala Leu Asn Ala Val Phe Val Glu Asn Phe Gln Arg Asp
210 215 220
Pro Thr Leu Thr Trp Gln Tyr Phe Gly Ser Ala Thr Gly Phe Phe Arg
225 230 235 240
Ile Tyr Pro Gly Ile Lys Trp Thr Pro Asp Glu Asn Gly Val Ile Thr
245 250 255
Phe Asp Cys Arg Asn Arg Gly Trp Tyr Ile Gln Ala Ala Thr Ser Pro
260 265 270
Lys Asp Ile Val Ile Leu Val Asp Val Ser Gly Ser Met Lys Gly Leu
275 280 285
Arg Met Thr Ile Ala Lys His Thr Ile Thr Thr Ile Leu Asp Thr Leu
290 295 300
Gly Glu Asn Asp Phe Ile Asn Ile Ile Ala Tyr Asn Asp Tyr Val His
305 310 315 320
Tyr Ile Glu Pro Cys Phe Lys Gly Ile Leu Val Gln Ala Asp Arg Asp
325 330 335
Asn Arg Glu His Phe Lys Leu Leu Val Glu Glu Leu Met Val Lys Gly
340 345 350
Val Gly Val Val Asp Gln Ala Leu Arg Glu Ala Phe Gln Ile Leu Lys
355 360 365
Gln Phe Gln Glu Ala Lys Gln Gly Ser Leu Cys Asn Gln Ala Ile Met
370 375 380
Leu Ile Ser Asp Gly Ala Val Glu Asp Tyr Glu Pro Val Phe Glu Lys
385 390 395 400
Tyr Asn Trp Pro Asp Cys Lys Val Arg Val Phe Thr Tyr Leu Ile Gly
405 410 415
Arg Glu Val Ser Phe Ala Asp Arg Met Lys Trp Ile Ala Cys Asn Asn
420 425 430
Lys Gly Tyr Tyr Thr Gln Ile Ser Thr Leu Ala Asp Thr Gln Glu Asn
435 440 445
Val Met Glu Tyr Leu His Val Leu Ser Arg Pro Met Val Ile Asn His
450 455 460
Asp His Asp Ile Ile Trp Thr Glu Ala Tyr Met Asp Ser Lys Leu Leu
465 470 475 480
Ser Ser Gln Ala Gln Ser Leu Thr Leu Leu Thr Thr Val Ala Met Pro
485 490 495
Val Phe Ser Lys Lys Asn Glu Thr Arg Ser His Gly Ile Leu Leu Gly
500 505 510
Val Val Gly Ser Asp Val Ala Leu Arg Glu Leu Met Lys Leu Ala Pro
515 520 525
Arg Tyr Lys Leu Gly Val His Gly Tyr Ala Phe Leu Asn Thr Asn Asn

530 535 540
Gly Tyr Ile Leu Ser His Pro Asp Leu Arg Pro Leu Tyr Arg Glu Gly
545 550 555 560
Lys Lys Leu Lys Pro Lys Pro Asn Tyr Asn Ser Val Asp Leu Ser Glu
565 570 575
Val Glu Trp Glu Asp Gln Ala Glu Ser Leu Arg Thr Ala Met Ile Asn
580 585 590
Arg Glu Thr Gly Thr Leu Ser Met Asp Val Lys Val Pro Met Asp Lys
595 600 605
Gly Lys Arg Val Leu Phe Leu Thr Asn Asp Tyr Phe Phe Thr Asp Ile
610 615 620
Ser Asp Thr Pro Phe Ser Leu Gly Val Val Leu Ser Arg Gly His Gly
625 630 635 640
Glu Tyr Ile Leu Leu Gly Asn Thr Ser Val Glu Glu Gly Leu His Asp
645 650 655
Leu Leu His Pro Asp Leu Ala Leu Ala Gly Asp Trp Ile Tyr Cys Ile
660 665 670
Thr Asp Ile Asp Pro Asp His Arg Lys Leu Ser Gln Leu Glu Ala Met
675 680 685
Ile Arg Phe Leu Thr Arg Lys Asp Pro Asp Leu Glu Cys Asp Glu Glu
690 695 700
Leu Val Arg Glu Val Leu Phe Asp Ala Val Val Thr Ala Pro Met Glu
705 710 715 720
Ala Tyr Trp Thr Ala Leu Ala Leu Asn Met Ser Glu Glu Ser Glu His
725 730 735
Val Val Asp Met Ala Phe Leu Gly Thr Arg Ala Gly Leu Leu Arg Ser
740 745 750
Ser Leu Phe Val Gly Ser Glu Lys Val Ser Asp Arg Lys Phe Leu Thr
755 760 765
Pro Glu Asp Glu Ala Ser Val Phe Thr Leu Asp Arg Phe Pro Leu Trp
770 775 780
Tyr Arg Gln Ala Ser Glu His Pro Ala Gly Ser Phe Val Phe Asn Leu
785 790 795 800
Arg Trp Ala Glu Gly Pro Glu Ser Ala Gly Glu Pro Met Val Val Thr
805 810 815
Ala Ser Thr Ala Val Ala Val Thr Val Asp Lys Arg Thr Ala Ile Ala
820 825 830
Ala Ala Ala Gly Val Gln Met Lys Leu Glu Phe Leu Gln Arg Lys Phe
835 840 845
Trp Ala Ala Thr Arg Gln Cys Ser Thr Val Asp Gly Pro Cys Thr Gln
850 855 860
Ser Cys Glu Asp Ser Asp Leu Asp Cys Phe Val Ile Asp Asn Asn Gly

865 870 875 880
Phe Ile Leu Ile Ser Lys Arg Ser Arg Glu Thr Gly Arg Phe Leu Gly
885 890 895
Glu Val Asp Gly Ala Val Leu Thr Gln Leu Leu Ser Met Gly Val Phe
900 905 910
Ser Gln Val Thr Met Tyr Asp Tyr Gln Ala Met Cys Lys Pro Ser Ser
915 920 925
His His His Ser Ala Ala Gln Pro Leu Val Ser Pro Ile Ser Ala Phe
930 935 940
Leu Thr Ala Thr Arg Trp Leu Leu Gln Glu Leu Val Leu Phe Leu Leu
945 950 955 960
Glu Trp Ser Val Trp Gly Ser Trp Tyr Asp Arg Gly Ala Glu Ala His
965 970 975
Lys His Lys Lys Gln Asp Pro Leu Gln Pro Cys Asp Thr Glu Tyr Pro
980 985 990
Val Phe Val Tyr Gln Pro Ala Ile Arg Glu Ala Asn Gly Ile Val Glu
995 1000 1005
Cys Gly Pro Cys Gln Lys Val Phe Val Val Gln Gln Ile Pro Asn Ser
1010 1015 1020
Asn Leu Leu Leu Leu Val Thr Asp Pro Thr Phe Cys Arg Met Gly Ser
1025 1030 1035 1040
Gly Pro Glu Ile Leu Thr Leu Thr Val Ala Ser Ala His Asn Ala Ser
1045 1050 1055
Val Lys Cys Asp Arg Met Arg Ser Gln Lys Leu Arg Arg Arg Pro Asp
1060 1065 1070
Ser Cys His Ala Phe His Pro Glu Glu Asn Ala Gln Asp Cys Gly Gly
1075 1080 1085
Ala Ser Asp Thr Ser Ala Ser Pro Pro Leu Leu Leu Leu Pro Val Cys
1090 1095 1100
Ala Trp Gly Leu Leu Pro Gln Leu Leu Arg
1105 1110

<210> 36
<211> 128
<212> PRT
<213> Homo sapiens

<400> 36
Met Ala Arg Ile Leu Leu Leu Phe Leu Pro Gly Leu Val Ala Val Cys
1 5 10 15
Ala Val His Gly Ile Phe Met Asp Arg Leu Ala Ser Lys Lys Leu Cys
20 25 30

Ala Asp Asp Glu Cys Val Tyr Thr Ile Ser Leu Ala Ser Ala Gln Glu
35 40 45

Asp Tyr Asn Ala Pro Asp Cys Arg Phe Ile Asn Val Lys Lys Gly Gln
50 55 60

Gln Ile Tyr Val Tyr Ser Lys Leu Val Lys Glu Asn Gly Ala Gly Glu
65 70 75 80

Phe Trp Ala Gly Ser Val Tyr Gly Asp Gly Gln Asp Glu Met Gly Val
85 90 95

Val Gly Tyr Phe Pro Arg Asn Leu Val Lys Glu Gln Arg Val Tyr Gln
100 105 110

Glu Ala Thr Lys Glu Val Pro Thr Thr Asp Ile Asp Phe Phe Cys Glu
115 120 125

<210> 37

<211> 215

<212> PRT

<213> Homo sapiens

<400> 37

Met Gly Leu Thr Trp Ile Leu Val Thr Ile Leu Leu Gly Gly Pro Gly
1 5 10 15

Val Gly Leu Pro Arg Ile Gln Gln Phe Phe Thr Ser Pro Glu Asn Ser
20 25 30

Val Thr Ala Glu Pro Arg Ala Arg Lys Tyr Lys Cys Gly Leu Pro Gln
35 40 45

Pro Cys Pro Glu Glu His Leu Ser Phe Arg Ile Val Ser Gly Ala Ala
50 55 60

Asn Val Ile Gly Pro Lys Ile Cys Leu Glu Asp Lys Met Leu Met Ser
65 70 75 80

Ser Val Lys Asp Asn Val Gly Arg Gly Leu Asn Ile Ala Leu Val Asn
85 90 95

Gly Val Ser Gly Glu Leu Leu Glu Ala Arg Ala Phe Asp Met Trp Ala
100 105 110

Gly Asp Val Asn Asp Leu Leu Lys Phe Ile Arg Pro Leu His Glu Gly
115 120 125

Thr Leu Val Phe Val Ala Ser Tyr Asp Asp Pro Ala Thr Lys Met Asn
130 135 140

Glu Glu Thr Arg Lys Leu Phe Ser Glu Leu Gly Ser Arg Asn Ala Lys
145 150 155 160

Asp Leu Ala Phe Arg Asp Ser Trp Val Phe Val Gly Ala Lys Gly Val
165 170 175

Gln Asn Lys Ser Pro Phe Glu Gln His Met Lys Asn Ser Lys His Thr

```
            180                 185                 190
Asn Lys Tyr Glu Gly Trp Pro Glu Ala Leu Glu Met Glu Gly Cys Ile
        195                 200                 205
Pro Arg Arg Ser Ile Ala Gly
    210                 215

      <210> 38
      <211> 230
      <212> PRT
      <213> Homo sapiens

      <400> 38
Met Arg Leu Ala Gly Pro Leu Arg Ile Val Ala Leu Ile Ile Ile Met
 1               5                  10                  15
Gly Leu Thr Trp Ile Leu Val Thr Ile Leu Leu Gly Gly Pro Gly Val
            20                  25                  30
Gly Leu Pro Arg Ile Gln Gln Phe Phe Thr Ser Pro Glu Asn Ser Val
        35                  40                  45
Thr Ala Glu Pro Arg Ala Arg Lys Tyr Lys Cys Gly Leu Pro Gln Pro
    50                  55                  60
Cys Pro Glu Glu His Leu Ser Phe Arg Ile Val Ser Gly Ala Ala Asn
65                  70                  75                  80
Val Ile Gly Pro Lys Ile Cys Leu Glu Asp Lys Met Leu Met Ser Ser
                85                  90                  95
Val Lys Asp Asn Val Gly Arg Gly Leu Asn Ile Ala Leu Val Asn Gly
            100                 105                 110
Val Ser Gly Glu Leu Leu Glu Ala Arg Ala Phe Asp Met Trp Ala Gly
        115                 120                 125
Asp Val Asn Asp Leu Leu Lys Phe Ile Arg Pro Leu His Glu Gly Thr
    130                 135                 140
Leu Val Phe Val Ala Ser Tyr Asp Asp Pro Ala Thr Lys Met Asn Glu
145                 150                 155                 160
Glu Thr Arg Lys Leu Phe Ser Glu Leu Gly Ser Arg Asn Ala Lys Asp
                165                 170                 175
Leu Ala Phe Arg Asp Ser Trp Val Phe Val Gly Ala Lys Gly Val Gln
            180                 185                 190
Asn Lys Ser Pro Phe Glu Gln His Met Lys Asn Ser Lys His Thr Asn
        195                 200                 205
Lys Tyr Glu Gly Trp Pro Glu Ala Leu Glu Met Glu Gly Cys Ile Pro
    210                 215                 220
Arg Arg Ser Ile Ala Gly
225                 230
```

<210> 39
<211> 436
<212> PRT
<213> Homo sapiens

<400> 39
Met Gln Gly Thr Pro Gly Gly Gly Thr Arg Pro Gly Pro Ser Pro Val
1 5 10 15
Asp Arg Arg Thr Leu Leu Val Phe Ser Phe Ile Leu Ala Ala Ala Leu
20 25 30
Gly Gln Met Asn Phe Thr Gly Asp Gln Val Leu Arg Val Leu Ala Lys
35 40 45
Asp Glu Lys Gln Leu Ser Leu Leu Gly Asp Leu Glu Gly Leu Lys Pro
50 55 60
Gln Lys Val Asp Phe Trp Arg Gly Pro Ala Arg Pro Ser Leu Pro Val
65 70 75 80
Asp Met Arg Val Pro Phe Ser Glu Leu Lys Asp Ile Lys Ala Tyr Leu
85 90 95
Glu Ser His Gly Leu Ala Tyr Ser Ile Met Ile Lys Asp Ile Gln Val
100 105 110
Leu Leu Asp Glu Glu Arg Gln Ala Met Ala Lys Ser Arg Arg Leu Glu
115 120 125
Arg Ser Thr Asn Ser Phe Ser Tyr Ser Ser Tyr His Thr Leu Glu Glu
130 135 140
Ile Tyr Ser Trp Ile Asp Asn Phe Val Met Glu His Ser Asp Ile Val
145 150 155 160
Ser Lys Ile Gln Ile Gly Asn Ser Phe Glu Asn Gln Ser Ile Leu Val
165 170 175
Leu Lys Phe Ser Thr Gly Gly Ser Arg His Pro Ala Ile Trp Ile Asp
180 185 190
Thr Gly Ile His Ser Arg Glu Trp Ile Thr His Ala Thr Gly Ile Trp
195 200 205
Thr Ala Asn Lys Ile Val Ser Asp Tyr Gly Lys Asp Arg Val Leu Thr
210 215 220
Asp Ile Leu Asn Ala Met Asp Ile Phe Ile Glu Leu Val Thr Asn Pro
225 230 235 240
Asp Gly Phe Ala Phe Thr His Ser Met Asn Arg Leu Trp Arg Lys Asn
245 250 255
Lys Ser Ile Arg Pro Gly Ile Phe Cys Ile Gly Val Asp Leu Asn Arg
260 265 270
Asn Trp Lys Ser Gly Phe Gly Gly Asn Gly Ser Asn Ser Asn Pro Cys

275 280 285

Ser Glu Thr Tyr His Gly Pro Ser Pro Gln Ser Glu Pro Glu Val Ala
290 295 300

Ala Ile Val Asn Phe Ile Thr Ala His Gly Asn Phe Lys Ala Leu Ile
305 310 315 320

Ser Ile His Ser Tyr Ser Gln Met Leu Met Tyr Pro Tyr Gly Arg Leu
325 330 335

Leu Glu Pro Val Ser Asn Gln Arg Glu Leu Tyr Asp Leu Ala Lys Asp
340 345 350

Ala Val Glu Ala Leu Tyr Lys Val His Gly Ile Glu Tyr Ile Phe Gly
355 360 365

Ser Ile Ser Thr Thr Leu Tyr Val Ala Ser Gly Ile Thr Val Asp Trp
370 375 380

Ala Tyr Asp Ser Gly Ile Lys Tyr Ala Phe Ser Phe Glu Leu Arg Asp
385 390 395 400

Thr Gly Gln Tyr Gly Phe Leu Leu Pro Ala Thr Gln Ile Ile Pro Thr
405 410 415

Ala Gln Glu Thr Trp Met Ala Leu Arg Thr Ile Met Glu His Thr Leu
420 425 430

Asn His Pro Tyr
435

<210> 40
<211> 419
<212> PRT
<213> Homo sapiens

<400> 40

Met Arg Thr Leu Leu Val Phe Ser Phe Ile Leu Ala Ala Ala Leu Gly
1 5 10 15

Gln Met Asn Phe Thr Gly Asp Gln Val Leu Arg Val Leu Ala Lys Asp
20 25 30

Glu Lys Gln Leu Ser Leu Leu Gly Asp Leu Glu Gly Leu Lys Pro Gln
35 40 45

Lys Val Asp Phe Trp Arg Gly Pro Ala Arg Pro Ser Leu Pro Val Asp
50 55 60

Met Arg Val Pro Phe Ser Glu Leu Lys Asp Ile Lys Ala Tyr Leu Glu
65 70 75 80

Ser His Gly Leu Ala Tyr Ser Ile Met Ile Lys Asp Ile Gln Val Leu
85 90 95

Leu Asp Glu Glu Arg Gln Ala Met Ala Lys Ser Arg Arg Leu Glu Arg
100 105 110

Ser Thr Asn Ser Phe Ser Tyr Ser Ser Tyr His Thr Leu Glu Glu Ile
115 120 125

Tyr Ser Trp Ile Asp Asn Phe Val Met Glu His Ser Asp Ile Val Ser
130 135 140

Lys Ile Gln Ile Gly Asn Ser Phe Glu Asn Gln Ser Ile Leu Val Leu
145 150 155 160

Lys Phe Ser Thr Gly Gly Ser Arg His Pro Ala Ile Trp Ile Asp Thr
165 170 175

Gly Ile His Ser Arg Glu Trp Ile Thr His Ala Thr Gly Ile Trp Thr
180 185 190

Ala Asn Lys Ile Val Ser Asp Tyr Gly Lys Asp Arg Val Leu Thr Asp
195 200 205

Ile Leu Asn Ala Met Asp Ile Phe Ile Glu Leu Val Thr Asn Pro Asp
210 215 220

Gly Phe Ala Phe Thr His Ser Met Asn Arg Leu Trp Arg Lys Asn Lys
225 230 235 240

Ser Ile Arg Pro Gly Ile Phe Cys Ile Gly Val Asp Leu Asn Arg Asn
245 250 255

Trp Lys Ser Gly Phe Gly Gly Asn Gly Ser Asn Ser Asn Pro Cys Ser
260 265 270

Glu Thr Tyr His Gly Pro Ser Pro Gln Ser Glu Pro Glu Val Ala Ala
275 280 285

Ile Val Asn Phe Ile Thr Ala His Gly Asn Phe Lys Ala Leu Ile Ser
290 295 300

Ile His Ser Tyr Ser Gln Met Leu Met Tyr Pro Tyr Gly Arg Leu Leu
305 310 315 320

Glu Pro Val Ser Asn Gln Arg Glu Leu Tyr Asp Leu Ala Lys Asp Ala
325 330 335

Val Glu Ala Leu Tyr Lys Val His Gly Ile Glu Tyr Ile Phe Gly Ser
340 345 350

Ile Ser Thr Thr Leu Tyr Val Ala Ser Gly Ile Thr Val Asp Trp Ala
355 360 365

Tyr Asp Ser Gly Ile Lys Tyr Ala Phe Ser Phe Glu Leu Arg Asp Thr
370 375 380

Gly Gln Tyr Gly Phe Leu Leu Pro Ala Thr Gln Ile Ile Pro Thr Ala
385 390 395 400

Gln Glu Thr Trp Met Ala Leu Arg Thr Ile Met Glu His Thr Leu Asn
405 410 415

His Pro Tyr

<210> 41

<211> 119
<212> PRT
<213> Homo sapiens

<400> 41

Met Trp Ser Leu Pro Pro Ser Arg Ala Leu Ser Cys Ala Pro Leu Leu
1 5 10 15
Leu Leu Phe Ser Phe Gln Phe Leu Val Thr Tyr Ala Trp Arg Phe Gln
20 25 30
Glu Glu Glu Glu Trp Asn Asp Gln Lys Gln Ile Ala Val Tyr Leu Pro
35 40 45
Pro Thr Leu Glu Phe Ala Val Tyr Thr Phe Asn Lys Gln Ser Lys Asp
50 55 60
Trp Tyr Ala Tyr Lys Leu Val Pro Val Leu Ala Ser Trp Lys Glu Gln
65 70 75 80
Val Asp Glu His Ile Leu Phe Cys Thr Ser Val Gln His Arg Leu Leu
85 90 95
Ser Asp Gly Gln Gly Trp Gln Arg Val Gly Gln Gly Leu Thr Arg Thr
100 105 110
Pro Gly Ser Pro Phe Val Val
115

<210> 42
<211> 148
<212> PRT
<213> Homo sapiens

<400> 42

Met Ser Ser Pro Gln Arg Arg Lys Ala Met Pro Trp Ala Leu Ser Leu
1 5 10 15
Leu Leu Met Gly Phe Gln Leu Leu Val Thr Tyr Ala Trp Cys Ser Glu
20 25 30
Glu Glu Met Gly Gly Asn Asn Lys Ile Val Gln Asp Pro Met Phe Leu
35 40 45
Ala Thr Val Glu Phe Ala Leu Asn Thr Phe Asn Val Gln Ser Lys Glu
50 55 60
Glu His Ala Tyr Arg Leu Leu Arg Val Leu Ser Ser Trp Arg Glu Asp
65 70 75 80
Ser Met Asp Arg Lys Met Val Phe Ser Met Asn Leu Gln Leu Arg Gln
85 90 95
Thr Val Cys Arg Lys Phe Glu Asp Asp Ile Asp Asn Cys Pro Phe Gln
100 105 110

Glu Ser Leu Glu Leu Asn Asn Thr Phe Thr Cys Phe Phe Thr Ile Ser
115 120 125
Thr Arg Pro Trp Met Thr Gln Phe Ser Leu Leu Asn Lys Thr Cys Leu
130 135 140
Glu Gly Phe His
145

<210> 43
<211> 898
<212> PRT
<213> Homo sapiens

<400> 43
Met Arg Ala Ala Leu Trp Thr Leu Gly Leu Gly Pro Leu Leu Leu Asn
1 5 10 15
Leu Trp Ala Val Pro Ile Gly Gly Pro Gly Ala Leu Arg Leu Ala Tyr
20 25 30
Arg His Ser Thr Cys Asp Gly Val Val Leu Val Arg His His Gly Ala
35 40 45
Trp Gly Tyr Val Cys Asn Gln Glu Trp Thr Leu Ala Glu Ala Ser Val
50 55 60
Val Cys Arg Gln Leu Gly Cys Gly Pro Ala Val Gly Ala Pro Lys Tyr
65 70 75 80
Val Pro Leu Pro Gly Glu Met Ala Gln Pro Trp Leu His Asn Val Ser
85 90 95
Cys Arg Gly Asn Glu Ser Ser Leu Trp Glu Cys Ser Leu Gly Ser Trp
100 105 110
Cys Gln Ser Pro Cys Pro His Ala Trp Val Val Val Ala Leu Cys Ser
115 120 125
Asn Gly Thr Phe Arg Glu Leu Arg Leu Val Lys Gly Arg Ser Pro Cys
130 135 140
Ala Gly Leu Pro Glu Ile Arg Asn Val Asn Gly Val Asp Arg Leu Cys
145 150 155 160
Val Leu His Val Glu Glu Ala Met Val Phe Cys Arg Glu Leu Gly Cys
165 170 175
Gly Pro Val Leu Gln Ala Pro Arg Arg Asp Val Gly Val Val Arg Lys
180 185 190
Tyr Leu Ala Cys Arg Gly Thr Glu Pro Thr Ile Arg Ser Cys Arg Leu
195 200 205
Asp Asn Asn Phe Arg Ser Gly Cys Asp Leu Arg Leu Asp Ala Glu Val
210 215 220
Val Cys Ser Gly His Thr Glu Ala Arg Leu Val Gly Gly Glu His Pro

```
225                 230                 235                 240
Cys Ala Gly Arg Leu Glu Val Thr Trp Gly Thr Val Cys Asp Ala Ala
                245                 250                 255
Leu Asp Leu Ala Thr Ala His Val Val Cys Arg Glu Leu Gln Cys Gly
            260                 265                 270
Ala Val Val Ser Thr Pro Glu Gly Ala Arg Phe Gly Arg Gly Ser Gly
        275                 280                 285
Pro Val Trp Thr Glu Ala Phe Arg Cys Ala Gly Asn Glu Ser Leu Leu
    290                 295                 300
Phe His Cys Pro Arg Gly Arg Gly Ser Gln Cys Gly His Gly His Asp
305                 310                 315                 320
Ala Gly Leu Arg Cys Ser Glu Phe Arg Met Val Asn Gly Ser Ser Ser
                325                 330                 335
Cys Glu Gly Arg Val Glu Phe Gln Val Gln Gly Ser Trp Ala Pro Leu
            340                 345                 350
Cys Ala Thr His Trp Asp Ile Ala Asp Ala Thr Val Leu Cys His Gln
        355                 360                 365
Leu Asn Cys Gly Asn Ala Val Ala Ala Pro Gly Gly Gly His Phe Gly
    370                 375                 380
Asp Gly Asp Ala Ala Ile Trp Pro Asp Ala Phe His Cys Glu Gly Thr
385                 390                 395                 400
Glu Ser Tyr Leu Trp Asn Cys Pro Val Ser Thr Leu Gly Ala Pro Ala
                405                 410                 415
Cys Ala Pro Gly Asn Thr Ala Ser Ala Val Cys Ser Gly Leu Ala His
            420                 425                 430
Ala Leu Arg Leu Arg Glu Gly Gln Ser Arg Cys Asp Gly Arg Val Glu
        435                 440                 445
Val Ser Leu Asp Gly Val Trp Gly Arg Val Leu Asp Asp Ala Trp Asp
    450                 455                 460
Leu Arg Gly Ala Gly Val Val Cys Arg Gln Leu Gly Cys Arg Gly Ala
465                 470                 475                 480
Gln Gln Ala Tyr Asp Ala Pro Ala Pro Ser Arg Gly Ser Val Gln Val
                485                 490                 495
Ala Leu Ser Arg Val Arg Cys Leu Gly Thr Glu Thr Arg Leu Thr Gln
            500                 505                 510
Cys Asn Val Ser Ala Thr Leu Gln Glu Pro Ala Gly Thr Ser Arg Asp
        515                 520                 525
Ala Gly Val Val Cys Ser Gly Glu Val Gly Thr Ala Ser Pro Met Ala
    530                 535                 540
Arg Arg His Gly Ile Pro Gly Ala Leu Thr Leu Ser Leu His Arg Glu
545                 550                 555                 560
Pro Gln Gly Ala Ala Gly Arg Gly Ala Gly Ala Leu His Gly Gly Ala
```

565 570 575
Trp Gly Thr Val Cys Asp Asp Ala Trp Asp Leu Arg Asp Ala His Val
580 585 590
Val Cys Arg Gln Leu Gly Cys Gly Arg Ala Leu Ser Ala Leu Gly Ala
595 600 605
Ala His Phe Gly Ala Gly Ala Gly Arg Ile Trp Leu Asp Glu Leu Gly
610 615 620
Cys Gln Gly His Glu Ser Ala Leu Trp Gln Cys Pro Ser Ala Gly Trp
625 630 635 640
Gly Arg His Asp Trp Arg His Lys Glu Asp Ala Gly Val Phe Cys Ser
645 650 655
Glu Ser Val Ala Leu Arg Leu Arg Gly Gly Thr Cys Cys Cys Ala Gly
660 665 670
Trp Leu Asp Val Phe Tyr Asn Gly Thr Trp Gly Ala Met Cys Ser Asn
675 680 685
Ala Leu Lys Asp Leu Ser Leu Ser Ile Ile Cys Lys Gln Leu Gly Cys
690 695 700
Gly Val Trp Gly Val Gly Leu Ala Gly Glu Gln Ala Leu Pro Leu Ala
705 710 715 720
Gly Thr Gly Thr Ala Trp Val Asp Asn Ile Glu Cys Arg Arg Leu Pro
725 730 735
Asn Ser Thr Leu Trp Gln Cys Pro Ser His Pro Trp His Pro His Ser
740 745 750
Cys Asp Leu Arg Glu Gln Val Trp Ile Thr Cys Ala Val Thr Ala Ala
755 760 765
Pro Phe Ala Glu Glu Gly Ala Leu Arg Val Arg Gly Gly Glu Asp Arg
770 775 780
Cys Ser Gly Arg Val Glu Leu Trp His Ala Gly Ser Trp Gly Thr Val
785 790 795 800
Cys Asp Asp Gly Trp Asp Leu Ala Asp Ala Glu Val Val Cys Arg Gln
805 810 815
Leu Gly Cys Gly Arg Ala Val Ala Ala Leu Gly Ala Ala Ala Phe Gly
820 825 830
Pro Gly Ser Gly Pro Val Trp Leu Asp Glu Val Gly Cys Arg Gly Ser
835 840 845
Glu Ala Ser Leu Trp Gly Cys Pro Ala Glu Arg Trp Gly Arg Gly Asp
850 855 860
Arg Ala His Glu Glu Asp Ala Gly Val Arg Cys Trp Gly Glu Trp Gly
865 870 875 880
Ala Val Gly Ser Arg Ser Trp Gly Arg Gln Arg Ala Leu Gly Trp Ser
885 890 895
Gln Ser

<210> 44
<211> 426
<212> PRT
<213> Homo sapiens

<400> 44
Met Ala Gly Leu Gly Phe Trp Gly His Pro Ala Gly Pro Leu Leu Leu
1 5 10 15
Leu Leu Leu Leu Val Leu Pro Pro Arg Ala Leu Pro Glu Gly Pro Leu
20 25 30
Val Phe Val Ala Leu Val Phe Arg His Gly Asp Arg Ala Pro Leu Ala
35 40 45
Ser Tyr Pro Met Asp Pro His Lys Glu Val Ala Ser Thr Leu Trp Pro
50 55 60
Arg Gly Leu Gly Gln Leu Thr Thr Glu Gly Val Arg Gln Gln Leu Glu
65 70 75 80
Leu Gly Arg Phe Leu Arg Ser Arg Tyr Glu Ala Phe Leu Ser Pro Glu
85 90 95
Tyr Arg Arg Glu Glu Val Tyr Ile Arg Ser Thr Asp Phe Asp Arg Thr
100 105 110
Leu Glu Ser Ala Gln Ala Asn Leu Ala Gly Leu Phe Pro Glu Ala Ala
115 120 125
Pro Gly Ser Pro Glu Ala Arg Trp Arg Pro Ile Pro Val His Thr Val
130 135 140
Pro Val Ala Glu Asp Lys Leu Leu Arg Phe Pro Met Arg Ser Cys Pro
145 150 155 160
Arg Tyr His Glu Leu Leu Arg Glu Ala Thr Glu Ala Ala Glu Tyr Gln
165 170 175
Glu Ala Leu Glu Gly Trp Thr Gly Phe Leu Ser Arg Leu Glu Asn Phe
180 185 190
Thr Gly Leu Ser Leu Val Gly Glu Pro Leu Arg Arg Ala Trp Lys Val
195 200 205
Leu Asp Thr Leu Met Cys Gln Gln Ala His Gly Leu Pro Leu Pro Ala
210 215 220
Trp Ala Ser Pro Asp Val Leu Arg Thr Leu Ala Gln Ile Ser Ala Leu
225 230 235 240
Asp Ile Gly Ala His Val Gly Pro Pro Arg Ala Ala Glu Lys Ala Gln
245 250 255
Leu Thr Gly Gly Ile Leu Leu Asn Ala Ile Leu Ala Asn Phe Ser Arg
260 265 270

Val Gln Arg Leu Gly Leu Pro Leu Lys Met Val Met Tyr Ser Ala His
275 280 285

Asp Ser Thr Leu Leu Ala Leu Gln Gly Ala Leu Gly Leu Tyr Asp Gly
290 295 300

His Thr Pro Pro Tyr Ala Ala Cys Leu Gly Phe Glu Phe Arg Lys His
305 310 315 320

Leu Gly Asn Pro Ala Lys Asp Gly Gly Asn Val Thr Val Ser Leu Phe
325 330 335

Tyr Arg Asn Asp Ser Ala His Leu Pro Leu Pro Leu Ser Leu Pro Gly
340 345 350

Cys Pro Ala Pro Cys Pro Leu Gly Arg Phe Tyr Gln Leu Thr Ala Pro
355 360 365

Ala Arg Pro Pro Ala His Gly Val Ser Cys His Gly Pro Tyr Glu Ala
370 375 380

Ala Ile Pro Pro Ala Pro Val Val Pro Leu Leu Ala Gly Ala Val Ala
385 390 395 400

Val Leu Val Ala Leu Ser Leu Gly Leu Gly Leu Leu Ala Trp Arg Pro
405 410 415

Gly Cys Leu Arg Ala Leu Gly Gly Pro Val
420 425

<210> 45
<211> 475
<212> PRT
<213> Homo sapiens

<400> 45

Met Leu Ala Ala Ser Ile Phe Arg Pro Thr Leu Leu Leu Cys Trp Leu
1 5 10 15

Ala Ala Pro Trp Pro Thr Gln Pro Glu Ser Leu Phe His Ser Arg Asp
20 25 30

Arg Ser Asp Leu Glu Pro Ser Pro Leu Arg Gln Ala Lys Pro Ile Ala
35 40 45

Asp Leu His Ala Ala Gln Arg Phe Leu Ser Arg Tyr Gly Trp Ser Gly
50 55 60

Val Trp Ala Ala Trp Gly Pro Ser Pro Glu Gly Pro Pro Glu Thr Pro
65 70 75 80

Lys Gly Ala Ala Leu Ala Glu Ala Val Arg Arg Phe Gln Arg Ala Asn
85 90 95

Ala Leu Pro Ala Ser Gly Glu Leu Asp Ala Ala Thr Leu Ala Ala Met
100 105 110

Asn Arg Pro Arg Cys Gly Val Pro Asp Met Arg Pro Pro Pro Pro Ser

115 120 125

Ala Pro Pro Ser Pro Pro Gly Pro Pro Pro Arg Ala Arg Ser Arg Arg

130 135 140

Ser Pro Arg Ala Pro Leu Ser Leu Ser Arg Arg Gly Trp Gln Pro Arg

145 150 155 160

Gly Tyr Pro Asp Gly Gly Ala Ala Gln Ala Phe Ser Lys Arg Thr Leu

165 170 175

Ser Trp Arg Leu Leu Gly Glu Ala Leu Ser Ser Gln Leu Ser Val Ala

180 185 190

Asp Gln Arg Arg Ile Val Ala Leu Ala Phe Arg Met Trp Ser Glu Val

195 200 205

Thr Pro Leu Asp Phe Arg Glu Asp Leu Ala Ala Pro Gly Ala Ala Val

210 215 220

Asp Ile Lys Leu Gly Phe Gly Arg Gly Ser Cys Glu Gly Ser Phe Asp

225 230 235 240

Thr Ala Phe Asp Trp Ile Arg Lys Glu Arg Asn Gln Tyr Gly Glu Val

245 250 255

Met Val Arg Phe Ser Thr Tyr Phe Phe Arg Asn Ser Trp Tyr Trp Leu

260 265 270

Tyr Glu Asn Arg Asn Asn Arg Thr Arg Tyr Gly Asp Pro Ile Gln Ile

275 280 285

Leu Thr Gly Trp Pro Gly Ile Pro Thr His Asn Ile Asp Ala Phe Val

290 295 300

His Ile Trp Thr Trp Lys Arg Asp Glu Arg Tyr Phe Phe Gln Gly Asn

305 310 315 320

Gln Tyr Trp Arg Tyr Asp Ser Asp Lys Asp Gln Ala Leu Thr Glu Asp

325 330 335

Glu Gln Gly Lys Ser Tyr Pro Lys Leu Ile Ser Glu Gly Phe Pro Gly

340 345 350

Ile Pro Ser Pro Leu Asp Thr Ala Phe Tyr Asp Arg Arg Gln Lys Leu

355 360 365

Ile Tyr Phe Phe Lys Glu Ser Leu Val Phe Ala Phe Asp Val Asn Arg

370 375 380

Asn Arg Val Leu Asn Ser Tyr Pro Lys Arg Ile Thr Glu Val Phe Pro

385 390 395 400

Ala Val Ile Pro Gln Asn His Pro Phe Arg Asn Ile Asp Ser Ala Tyr

405 410 415

Tyr Ser Tyr Ala Tyr Asn Ser Ile Phe Phe Phe Lys Gly Asn Ala Tyr

420 425 430

Trp Lys Val Val Asn Asp Lys Asp Lys Gln Gln Asn Ser Trp Leu Pro

435 440 445

Ala Asn Gly Leu Phe Pro Lys Lys Phe Ile Ser Glu Lys Trp Phe Asp

450 455 460
Val Cys Asp Val His Ile Ser Thr Leu Asn Met
465 470 475

<210> 46
<211> 529
<212> PRT
<213> Homo sapiens

<400> 46
Met Leu Ala Ala Ser Ile Phe Arg Pro Thr Leu Leu Leu Cys Trp Leu
1 5 10 15
Ala Ala Pro Trp Pro Thr Gln Pro Glu Ser Leu Phe His Ser Arg Asp
20 25 30
Arg Ser Asp Leu Glu Pro Ser Pro Leu Arg Gln Ala Lys Pro Ile Ala
35 40 45
Asp Leu His Ala Ala Gln Arg Phe Leu Ser Arg Tyr Gly Trp Ser Gly
50 55 60
Val Trp Ala Ala Trp Gly Pro Ser Pro Glu Gly Pro Pro Glu Thr Pro
65 70 75 80
Lys Gly Ala Ala Leu Ala Glu Ala Val Arg Arg Phe Gln Arg Ala Asn
85 90 95
Ala Leu Pro Ala Ser Gly Glu Leu Asp Ala Ala Thr Leu Ala Ala Met
100 105 110
Asn Arg Pro Arg Cys Gly Pro Arg Gly Tyr Pro Asp Gly Gly Ala Ala
115 120 125
Gln Ala Phe Ser Lys Arg Thr Leu Ser Trp Arg Leu Leu Gly Glu Ala
130 135 140
Leu Ser Ser Gln Leu Ser Val Ala Asp Gln Arg Arg Ile Val Ala Leu
145 150 155 160
Ala Phe Arg Met Trp Ser Glu Val Thr Pro Leu Asp Phe Arg Glu Asp
165 170 175
Leu Ala Ala Pro Gly Ala Ala Val Asp Ile Lys Leu Gly Phe Gly Arg
180 185 190
Gly Arg His Leu Gly Cys Pro Arg Ala Phe Asp Gly Ser Gly Gln Glu
195 200 205
Phe Ala His Ala Trp Arg Leu Gly Asp Ile His Phe Asp Asp Asp Glu
210 215 220
His Phe Thr Pro Pro Thr Ser Asp Thr Gly Ile Ser Leu Leu Lys Val
225 230 235 240
Ala Val His Glu Ile Gly His Val Leu Gly Leu Pro His Thr Tyr Arg
245 250 255

Thr Gly Ser Ile Met Gln Pro Asn Tyr Ile Pro Gln Glu Pro Ala Phe
260 265 270

Glu Leu Asp Trp Ser Asp Arg Lys Ala Ile Gln Lys Leu Tyr Gly Ser
275 280 285

Cys Glu Gly Ser Phe Asp Thr Ala Phe Asp Trp Ile Arg Lys Glu Arg
290 295 300

Asn Gln Tyr Gly Glu Val Met Val Arg Phe Ser Thr Tyr Phe Phe Arg
305 310 315 320

Asn Ser Trp Tyr Trp Leu Tyr Glu Asn Arg Asn Asn Arg Thr Arg Tyr
325 330 335

Gly Asp Pro Ile Gln Ile Leu Thr Gly Trp Pro Gly Ile Pro Thr His
340 345 350

Asn Ile Asp Ala Phe Val His Ile Trp Thr Trp Lys Arg Asp Glu Arg
355 360 365

Tyr Phe Phe Gln Gly Asn Gln Tyr Trp Arg Tyr Asp Ser Asp Lys Asp
370 375 380

Gln Ala Leu Thr Glu Asp Glu Gln Gly Lys Ser Tyr Pro Lys Leu Ile
385 390 395 400

Ser Glu Gly Phe Pro Gly Ile Pro Ser Pro Leu Asp Thr Ala Phe Tyr
405 410 415

Asp Arg Arg Gln Lys Leu Ile Tyr Phe Phe Lys Glu Ser Leu Val Phe
420 425 430

Ala Phe Asp Val Asn Arg Asn Arg Val Leu Asn Ser Tyr Pro Lys Arg
435 440 445

Ile Thr Glu Val Phe Pro Ala Val Ile Pro Gln Asn His Pro Phe Arg
450 455 460

Asn Ile Asp Ser Ala Tyr Tyr Ser Tyr Ala Tyr Asn Ser Ile Phe Phe
465 470 475 480

Phe Lys Gly Asn Ala Tyr Trp Lys Val Val Asn Asp Lys Asp Lys Gln
485 490 495

Gln Asn Ser Trp Leu Pro Ala Asn Gly Leu Phe Pro Lys Lys Phe Ile
500 505 510

Ser Glu Lys Trp Phe Asp Val Cys Asp Val His Ile Ser Thr Leu Asn
515 520 525

Met

<210> 47
<211> 402
<212> PRT
<213> Homo sapiens

<400> 47

```
Met Val Cys Ala Arg Ala Ala Leu Gly Pro Gly Ala Leu Trp Ala Ala
1               5                   10                  15
Ala Trp Gly Val Leu Leu Leu Thr Ala Pro Ala Gly Ala Gln Arg Gly
            20                  25                  30
Arg Lys Lys Val Val His Val Leu Glu Gly Glu Ser Gly Ser Val Val
        35                  40                  45
Val Gln Thr Ala Pro Gly Gln Val Val Ser His Arg Gly Gly Thr Ile
    50                  55                  60
Val Leu Pro Cys Arg Tyr His Tyr Glu Ala Ala Ala His Gly His Asp
65                  70                  75                  80
Gly Val Arg Leu Lys Trp Thr Lys Val Val Asp Pro Leu Ala Phe Thr
                85                  90                  95
Asp Val Phe Val Ala Leu Gly Pro Gln His Arg Ala Phe Gly Ser Tyr
            100                 105                 110
Arg Gly Arg Ala Glu Leu Gln Gly Asp Gly Pro Gly Asp Ala Ser Leu
        115                 120                 125
Val Leu Arg Asn Val Thr Leu Gln Asp Tyr Gly Arg Tyr Glu Cys Glu
    130                 135                 140
Val Thr Asn Glu Leu Glu Asp Asp Ala Gly Met Val Lys Leu Asp Leu
145                 150                 155                 160
Glu Gly Val Val Phe Pro Tyr His Pro Arg Gly Gly Arg Tyr Lys Leu
                165                 170                 175
Thr Phe Ala Glu Ala Gln Arg Ala Cys Ala Glu Gln Asp Gly Ile Leu
            180                 185                 190
Ala Ser Ala Glu Gln Leu His Ala Ala Trp Arg Asp Gly Leu Asp Trp
        195                 200                 205
Cys Asn Ala Gly Trp Leu Arg Asp Gly Ser Val Gln Tyr Pro Val Asn
    210                 215                 220
Arg Pro Arg Glu Pro Cys Gly Gly Leu Gly Gly Thr Gly Ser Ala Gly
225                 230                 235                 240
Gly Gly Gly Asp Ala Asn Gly Gly Leu Arg Asn Tyr Gly Tyr Arg His
                245                 250                 255
Asn Ala Glu Glu Arg Tyr Asp Ala Phe Cys Phe Thr Ser Asn Leu Pro
            260                 265                 270
Gly Arg Val Phe Phe Leu Lys Pro Leu Arg Pro Val Pro Phe Ser Gly
        275                 280                 285
Ala Ala Arg Ala Cys Ala Ala Arg Gly Ala Ala Val Ala Lys Val Gly
    290                 295                 300
Gln Leu Phe Ala Ala Trp Lys Leu Gln Leu Leu Asp Arg Cys Thr Ala
305                 310                 315                 320
Gly Trp Leu Ala Asp Gly Ser Ala Arg Tyr Pro Ile Val Asn Pro Arg
```

325 330 335
Ala Arg Cys Gly Gly Arg Arg Pro Gly Val Arg Ser Leu Gly Phe Pro
340 345 350
Asp Ala Thr Arg Arg Leu Phe Gly Val Tyr Cys Tyr Arg Ala Pro Gly
355 360 365
Ala Pro Asp Pro Ala Pro Gly Gly Trp Gly Trp Gly Trp Ala Gly Gly
370 375 380
Gly Gly Trp Ala Gly Gly Ala Arg Asp Pro Ala Ala Trp Thr Pro Leu
385 390 395 400
His Val

<210> 48
<211> 441
<212> PRT
<213> Homo sapiens

<400> 48
Met Leu Pro Ala Arg Cys Ala Arg Leu Leu Thr Pro His Leu Leu Leu
1 5 10 15
Val Leu Val Gln Leu Ser Pro Ala Arg Gly His Arg Thr Thr Gly Pro
20 25 30
Arg Phe Leu Ile Ser Asp Arg Asp Pro Gln Cys Asn Leu His Cys Ser
35 40 45
Arg Thr Gln Pro Lys Pro Ile Cys Ala Ser Asp Gly Arg Ser Tyr Glu
50 55 60
Ser Met Cys Glu Tyr Gln Arg Ala Lys Cys Arg Asp Pro Thr Leu Gly
65 70 75 80
Val Val His Arg Gly Arg Cys Lys Asp Ala Gly Gln Ser Lys Cys Arg
85 90 95
Leu Glu Arg Ala Gln Ala Leu Glu Gln Ala Lys Lys Pro Gln Glu Ala
100 105 110
Val Phe Val Pro Glu Cys Gly Glu Asp Gly Ser Phe Thr Gln Val Gln
115 120 125
Cys His Thr Tyr Thr Gly Tyr Cys Trp Cys Val Thr Pro Asp Gly Lys
130 135 140
Pro Ile Ser Gly Ser Ser Val Gln Asn Lys Thr Pro Val Cys Ser Gly
145 150 155 160
Ser Val Thr Asp Lys Pro Leu Ser Gln Gly Asn Ser Gly Arg Lys Asp
165 170 175
Asp Gly Ser Lys Pro Thr Pro Thr Met Glu Thr Gln Pro Val Phe Asp
180 185 190

Gly Asp Glu Ile Thr Ala Pro Thr Leu Trp Ile Lys His Leu Val Ile
195 200 205
Lys Asp Ser Lys Leu Asn Asn Thr Asn Ile Arg Asn Ser Glu Lys Val
210 215 220
Tyr Ser Cys Asp Gln Glu Arg Gln Ser Ala Leu Glu Glu Ala Gln Gln
225 230 235 240
Asn Pro Arg Glu Gly Ile Val Ile Pro Glu Cys Ala Pro Gly Gly Leu
245 250 255
Tyr Lys Pro Val Gln Cys His Gln Ser Thr Gly Tyr Cys Trp Cys Val
260 265 270
Leu Val Asp Thr Gly Arg Pro Leu Pro Gly Thr Ser Thr Arg Tyr Val
275 280 285
Met Pro Ser Cys Glu Ser Asp Ala Arg Ala Lys Thr Thr Glu Ala Asp
290 295 300
Asp Pro Phe Lys Asp Arg Glu Leu Pro Gly Cys Pro Glu Gly Lys Lys
305 310 315 320
Met Glu Phe Ile Thr Ser Leu Leu Asp Ala Leu Thr Thr Asp Met Val
325 330 335
Gln Ala Ile Asn Ser Ala Ala Pro Thr Gly Gly Gly Arg Phe Ser Glu
340 345 350
Pro Asp Pro Ser His Thr Leu Glu Glu Arg Val Val His Trp Tyr Phe
355 360 365
Ser Gln Leu Asp Ser Asn Ser Ser Asn Asp Ile Asn Lys Arg Glu Met
370 375 380
Lys Pro Phe Lys Arg Tyr Val Lys Lys Lys Ala Lys Pro Lys Lys Cys
385 390 395 400
Ala Arg Arg Phe Thr Asp Tyr Cys Asp Leu Asn Lys Asp Lys Val Ile
405 410 415
Ser Leu Pro Glu Leu Lys Gly Cys Leu Gly Val Ser Lys Glu Gly Gly
420 425 430
Ser Leu Gly Ser Phe Pro Gln Ala Lys
435 440

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/11797

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :C12N 5/10, 15/12, 15/63, 15/64; C07K 14/435, 14/47
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 435/69.1, 471, 71.1, 71.2, 471, 252.3, 254.11, 325, 320 1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 92/05256 A1 (GENETICS INSTITUTE, INC.) 02 April 1992 (02.04.92), see entire document.	1-7

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
02 JULY 2001

Date of mailing of the international search report
06 AUG 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
Bridgers for
PREMA MERTZ
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/11797

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

530/350; 435/69.1, 471, 71.1, 71.2, 471, 252.3, 254.11, 325, 320.1